VOL. IX JUNE, 1925 No. 2

The Journal of APPLIED PSYCHOLOGY

EDITORS

JAMES P. PORTER
Ohio University

WILLIAM F. BOOK
Indiana University

WITH A

BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

SARA M. STEINCHEIFER, PH.D. The Speech of Five Hundred Freshmen College Women 109

H. L. HOLLINGWORTH. When is a Man Intoxicated? 122

MORRIS S. VITELES. The Clinical Viewpoint in Vocational Selection 131

GEORGE C. BRANDENBURG. Analyzing Personality. Part I. 139

GRACE H. MANSON. Group Differences in Intelligence Tests. The Relative Difficulty of Types of Questions 156

H. K. NIXON. Two Studies of Attention to Advertisements 176

PERCIVAL M. SYMONDS. Notes on Rating 188

WINIFRED F. HYDE AND J. P. GUILFORD. A Test for Classification of Students in Chemistry 196

NOTES AND NEWS 203

BOOK REVIEWS 205

NEW BOOKS AND PAMPHLETS RECEIVED 212

PUBLISHED BY THE EDITORS
THE WILLIAMS & WILKINS CO.
Mount Royal and Guilford Avenues
Baltimore, Md.

Entered as second class matter June 2, 1925, at the post office at Baltimore, Maryland, under the act of March 3, 1879.

Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917, Authorized December 31, 1923.

COPYRIGHT 1925, BY JAMES P. PORTER

The Journal of Applied Psychology

CO-OPERATING EDITORS

J. W. V. Benedek, Carnegie Institute of Technology; James E. Downey, University of Wyoming; Rossee H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Jastrow Jastrow, University of Wisconsin; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College, Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Harvard Graduate School of Business Administration; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. B. Wallace Wallin, Teachers College, Miami University; G. M. Whipple, University of Michigan; Helen Thompson Woolley, Merrill-Palmer School, Detroit.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, salesmanship, public speaking, industrial and commercial work, and problems of business appeal. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal: in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published quarterly. The subscription price is $6.00 a year. Single numbers $1.50. Remittances and business communications should be addressed to the Editors, James P. Porter, Ohio University, Athens, Ohio, or to William F. Book, Indiana University, Bloomington, Indiana.

THE SPEECH OF FIVE HUNDRED FRESHMAN COLLEGE WOMEN[1]

SARA M. STINCHFIELD, PH.D.

Mount Holyoke College

Since the fall of 1922, standardized tests in voice and speech have been given to all entering students at Mount Holyoke College. The chief value of such tests is the ease with which it becomes possible to classify students for training. Students handicapped by speech and personality difficulties may be selected for training in small groups or individually. Students who seem to conform to a satisfactory standard may be excused from speech requirements; superior students in speech attainment may be discovered at the outset and recommended to do special work or to elect suitable speech courses.

Previous experience in various colleges seems to indicate that unless students are subjected to some examination in speech, many are graduated without having made any material progress in overcoming provincialisms, speech handicaps, personality difficulties, and even vulgarities of expression. Speech departments frequently never hear of such students. Many serious cases of lisping, stuttering, and indistinct utterance are present in every college, and personality difficulties which interfere seriously with college social adjustments are often ignored. As yet, the majority of our colleges are doing little to remove these handicaps.

There should be no cause for disagreement along different lines of speech work. There seems to have been a certain amount of misunderstanding, however, in the past because of the failure to recognize the dividing line between the artistic

[1] Paper given at the joint conference of the New England and Eastern Teachers of Speech, at Harvard, April, 1924.

and scientific approach to speech problems. We need to know how to harmonize the two.

In music much the same condition exists—there is the difference between teaching a child to play the piano and teaching him to tune the instrument. Most people realize when something is wrong with the piano, and they send for the expert, the piano-tuner. I have heard of a musician who found an old square piano quite out of tune, and as the owner was willing to risk his tampering with it, he undertook to tune it. Having a good ear he made it sound very well in the key of C in which he did the tuning, but not understanding the principle of the tempered scale, he left the piano out of tune in every key involving the use of sharps and flats.

The same thing often occurs in the attempt to correct minor speech defects. A person trained along interpretive lines, like the pianist, may modify poise, appearance and speech reactions of a speaker so that under controlled conditions or in some situation the speech is correct and even excellent, but various faulty temperamental difficulties in the individual have not been even suspected. Many types of speech defects with "introversion" or emotional imbalance need the scientific approach for their treatment, as does the successful tuning of the piano, otherwise one may obtain an accentuation, rather than a correction of the difficulty. For instance, in stuttering a person is often referred to various unscientific schools which "guarantee a cure." Of course no reputable physician ever *guarantees* a cure. I have just received a letter from a lad protesting strongly against such unethical methods. He writes: "Is there any standard method for the correction of stammering, which is at least partly reliable? There are so many relapses, after one has taken various *guaranteed cures*, that one loses hope!"

The modern scientific methods of correcting speech disorders have followed the principles of physical medical treatment, i.e., not to treat *symptoms*, but *causes*. Take rheumatism for instance—our grandparents rubbed oil on lame joints to secure relief from pain, but there was no known *cure* for rheu-

matism. Nowadays we find substances within the joints which produce rheumatism and the causes are often found in autointoxication produced by such conditions as diseased tonsils, decayed teeth, etc. Physicians do not treat *headaches*, though there are plenty of patent medicines which claim to do so. The scientific medical man seeks the cause of the headache and attempts to remove the symptoms by changing the causal conditions.

The purpose of our speech examination is selective; the plan being that students handicapped by ineffective voice and speech or personality difficulties and maladjustments may receive the necessary training; that students who conform to a satisfactory standard may be excused from training; and that special talent may be encouraged.

To secure a speech index or speech rating, the standardized Blanton-Stinchfield Speech Measurements (6) were used, which include an estimate of personality, of articulation of all consonant and vowel sounds in the language, of speech and voice, in the form of a series of objective and subjective measurements. The report of our findings is largely in tabular form with a brief explanation of each table.

Table 1 gives the distribution by groups of the entering classes of 1922 and 1923, with percentage of students in each group.

The striking fact in this study is, that although about one-third of the students last year and nearly half of the students this year passed the speech requirement, at least one-sixth of the Freshman class need definite corrective work, individually or in small groups. This makes a group of more than fifty in this year's class.

Besides these another considerable portion of the class, from one-fifth to one-fourth need speech training which is postponed to the Sophomore year. This group consists of less urgent cases than those in the corrective group. There seems to be no doubt that they should receive training before being allowed to graduate.

It is a source of great satisfaction however to find an equal

number (one-fifth to one-fourth), who seem to have sufficient native ability to justify us in encouraging them to develop

TABLE 1

GROUP		YEAR 1922		YEAR 1923	
		Number of students	Per cent of class	Number of students	Per cent of class
			per cent		*per cent*
I	Required Freshman work in speech correction	39	18	53	16
II	Required Sophomore speech work (ineffective speech)	42	19	80	24
III	Excused from speech requirement (average group)	75	35	144	42
IV	Recommended to elect speech because of good standing in the tests	60	28	63	18
Totals		216	100	340	100

Total number of students for both years, 556.

TABLE 2

Frequency table showing types of speech difficulties for which students are placed in Groups I and II (required Freshman and Sophomore speech classes), totalling 133 students. 1923–24

	PER CENT
Oral inactivity and ineffective speech	41.0
Lisping (on s and z sounds)	17.0
Faults of voice, rhythm and breath control	25.0
Stuttering, hesitation or other indications of nervousness	6.0
Foreign accent	2.3
Letter substitution	0.7
Mixed (more than one type of speech difficulty)	8.0
Total	100.0

their talents by elective work, as voice, personality and social adjustment in these cases seem to be of the normal, well-balanced type.

The total number of entering students, September, 1923, was 340. Of these 16 per cent were classified in required Freshman corrective group, and 24 per cent were classified in required Sophomore speech classes (see table 1).

The largest group in this table consists of those who show general oral inactivity and ineffective speech, defects which seem to be based upon the general constitutional make-up of the individual. These cases have been particularly difficult for teachers in speech classes, being reported as "uninteresting" and "unreceptive" or "unresponsive" to training.

Speech disturbances disappear when the individual feels free to express himself; that is, when he can give and take emotions in a normal way. One must remember that the neurotic symptoms are, as it were, a morbid gain for the patient, and unconsciously he refuses to give them up. Chronic stutterers, like other chronic neurotics, make use of their stuttering to escape from all kinds of difficulties. Stutterers, the world over, stutter most when they are confronted with disagreeable tasks, and a great many of them who have remained "cured" for years, relapse when confronted with a painful situation. So that our treatment consists in showing the patient how to face reality. There is no royal road to the treatment of speech disturbances. It is the most difficult neurosis that one has to manage. There are very few *cures* effected among the adult chronic stutterers (1).

All motor tics and habit spasms have an unfortunate effect on speech, as the motor activities tend to become less rhythmical and induce a condition of nervous exhaustion. Certain tics because they interfere directly with the free use of the vocal and speech mechanisms, must be specially considered by the speech worker (2).

Many of these tics have their foundation in some slight physical disability, which should be removed. Some, unfortunately, have their foundation in psychic disability, the removal of which is a longer and more difficult process. All, whether of physical or of psychic origin, have a definite psychic value. *The organism does not continue useless movements.* They may be apparently useless, may even seem so to the person who has acquired them, but that they are emblems of some need or serve some purpose in the emotional life of the person cannot be doubted (3).

Although the experiment is relatively new at Mount Holyoke, we feel justified in saying that the more severe types of speech

defects, emotional imbalance and lisping, sometimes give the better prognosis, because in such cases the student more readily recognizes her handicap, and exerts herself to overcome it. Hence, what sometimes seems an utterly hopeless case, may turn out to be most hopeful, while comparatively mild cases, may make little if any improvement. This applies particularly to hysterical stutterers, in whom surprisingly good results are often obtained in a short time, while seemingly mild stutterers show little improvement after the expenditure of much time and effort on the part of the teacher and the pupil. This does not hold for chronic stutterers, or for many cases of speech disorder where there are severe organic and functional factors in the background, such as oral and nasal malformations, epilepsy, infantile paralysis, etc.

It will doubtless be surprising to many persons to learn that among the highly selected group of girls who are admitted to college, we find such a variety of types of speech difficulties, all of which are recognized as more or less serious handicaps in post-college days. Considering the great stress that is placed upon appearance, address and behavior by employers of women, it seems a very great misfortune for our girls to fail of success because of a remedial speech difficulty.

A striking illustration of the need of speech training during college days came to the attention of the writer recently, when a high school teacher of English was referred to us. Her speech was of the oral inactivity type in which the tongue contacts for various consonants were inaccurately made, and there was a disagreeable lateral lisp present. No one had ever called the attention of the girl or her parents to the presence of this speech defect. It was only when she began to teach that she realized that she possessed a speech handicap.

Another instance came to our attention recently in consultation with a dentist who has suffered a lifetime from feelings of inferiority due to congenital cleft palate. This condition was allowed to go uncorrected for several years with the result that the youth became embittered, sensitive, brooding and morose to the extent that in adult manhood he cannot throw off these

moody tendencies, even though the condition which originally produced them, has been removed. He recently refused to appear before two distinguished dental organizations, where he had been invited to demonstrate his newly invented dental appliance—his refusal being based on fear of failure or ridicule, even though he now has a pleasing personality and a rich resonant voice quite free from nasality.

If the speech defects are in reality symptoms of an unwholesome temperamental condition, attempts to cure the symptoms without reference to the temperamental factors are found to be ineffective. It is obvious that recurrence of the symptoms is to be expected. Occasionally the speech trainer may be fortunate enough to *resolve the conflict*, and so bring about permanent relief. Temporary alleviation which relapses as soon as the patient dispenses with treatment indicates how great is the force of suggestion, *how largely the speech defective depends upon others for the stamina necessary for proper speech* (4).

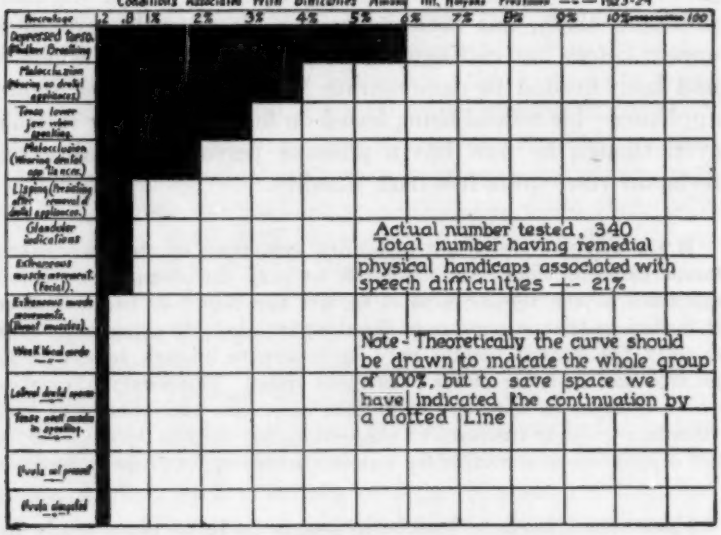
Moreover adequate provision seems to have been made for corrective work for most of the physical defects which do appear. Access to the health and medical reports of these students was possible through the coöperation of the department of physical education and it seems clear that very few students with serious physical defects are admitted.

There is a striking difference between the findings given in graph I and graph II, in which mental and physical difficulties are indicated.

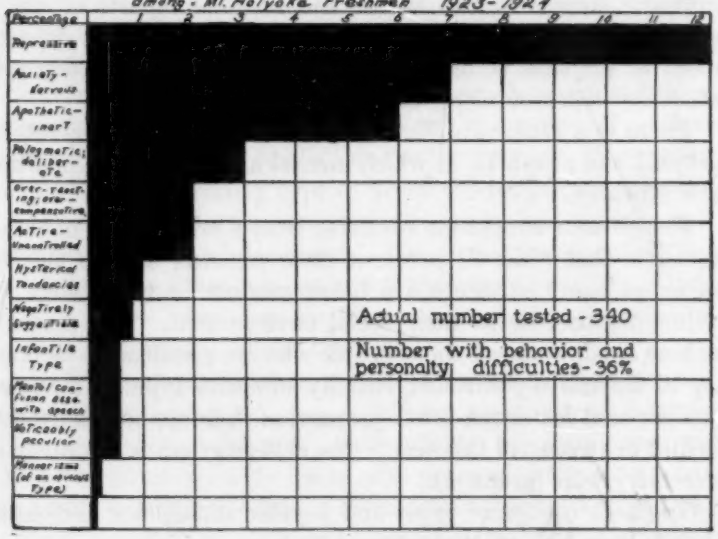
The obvious conclusion resulting from a comparison of these tables is that while 21 per cent show physical defects, nearly twice as many students are below normal in the ability to adjust themselves to their social environment. (By normal, as here used, we mean those people who are classified as belonging to the active-controlled, socially adjustive type.) Twenty-one per cent is a much lower proportion than one would expect to find in a group of this size. The college group is obviously a selected group (graph I).

Graph II (behavior types and psychic difficulties) indicates that 36 per cent need attention.

Graph I — Showing Physical Conditions Associated With Difficulties Among Mt. Holyoke Freshmen — 1923-24


Percentage
.2 .8 1% 2% 3% 4% 5% 6% 7% 8% 9% 10% 100
Depressed Facial Bladder Breathing
Malocclusion (showing no dental appliances)
Tense lower jaw when speaking
Malocclusion (showing dental appliances)
Lisping (persisting after removal of dental appliances)
Glandular indications
Extraneous muscle movements (facial).
Extraneous muscle movements (throat muscles).
Weak Nasal sounds
Labial dental sounds
Tense vocal cords in speaking.
Uvula relaxed
Uvula elongated
Actual number tested, 340
Total number having remedial physical handicaps associated with speech difficulties — 21%
Note - Theoretically the curve should be drawn to indicate the whole group of 100%, but to save space we have indicated the continuation by a dotted Line

Graph II showing Behavior Types and Psychic Difficulties among Mt. Holyoke Freshmen 1923-1924



Adjustment to the exigencies of life is the great and fundamental mental law on which all human intercourse is based, and all human happiness. The child cannot dominate the world. If he could it would not bring him happiness, because it would not bring him friends. The earlier he is taught the simple practice of looking at his needs from the point of view of society, desiring for himself only those things that would be reasonable if demanded by the whole race, and accepting those things that it were not well should be changed, the greater will be his opportunity for normal emotional life, and healthful spiritual development. Normal speech is only present where these two attributes are present (3).

In the Mental Hygiene Primer, Dr. Pratt writes:

Although men and women in years, many persons because of their mental make-up, for which early parental influence and environment are chiefly to blame, habitually carry over into adult life many of the dependent habits and customs of thinking of their childhood, only to find that reality demands a life adjustment based on mature mechanisms.

The whole mental hygiene movement is rapidly getting back to its starting point—early childhood—and is dedicating a large portion of its program to this period. During this flexible and formative age sound habits of mental health can best be instilled and faulty ones corrected.

Mental hygienists are stressing one great point, namely, that in most cases of nervousness, in many cases of delinquency, in some cases of insanity and in almost all cases of child behavior or conduct disorder, the trail leads inevitably and directly back to the *home* and the *parents*. And this fact operates in just the same fashion and with almost as much vigor and frequency among families of the well-to-do as it does in the tenements (5).

Social and economic maladjustments, behavior problems, psychopathic reactions ranging from feelings of inferiority to suicidal tendencies, often have their roots in mental difficulties of individuals who may influence large groups. A peculiar mental "twist," warped judgment, or feelings of inferiority due to repeated failure, may be the explanation of incipient and mild behavior disorders among college girls, particularly those of a repressive, introvert type. While there is much to be said in favor of a limited amount of "New England reserve," there are undoubtedly cases of anxiety neurosis which result from

severe repression of the normal play life, or from severe criticism by the elders of comparatively mild offenses in adolescence. Graph II suggests rather a large per cent of such cases among our entering college girls.

We sometimes find an "apathetic-inert" girl, who is passive rather than active in her attitude toward life; or she may be of a sophisticated, "poseur" type Her reactions seem characteristic of a physically passive, rather than of an active individual—not unlike the frail, exotic Burne-Jones ideal of half a century ago This is not the type of girl who wins prizes in the athletic contests of today. She may be a good student, and yet suffer from innumerable minor "ailments."

The "phlegmatic or deliberative" type differs from the above in mental and in physical characteristics. This type of student inclines to be faithful and plodding, but rarely takes high honors. Instead of being erratic, she is steady, industrious and fairly regular in habits, but less artistic, more easy-going and less easily disturbed than the former type. She may not fail in her work but she is inclined to be slow-moving, lethargic, deliberate or procrastinating.

The "over-reacting or over-compensative" type often results from the desire to conceal some real or fancied inferiority. The girl is anxious to be well-liked, and being unable to attract friends easily enough in ordinary ways, adopts bizarre, fantastic or clownish tricks and mannerisms which make an appeal to her group, because of their unusual character, novelty and seeming originality.

There are always a few girls who find their way to college, who have been singularly isolated from playmates of their own age, or who have developed a feeling of social inferiority as a result of some real or imagined family stigma. The impression made upon the memory of the growing girl seems well-nigh irradicable. Thus, the daughter of a divorcee, in one instance, appeared to be a sophisticated "flapper" of the cosmopolitan type, world-weary at eighteen, "bored" by most of the youths of her acquaintance. In reality she was a girl of unusual mental endowments, who stood high in college work, being

witty and attractive socially, despite her affectation of "boredom." She was distrustful of men and wary of their attentions, because of her knowledge of her mother's unhappy marriage. When some of her unethical reactions were discussed with her by an older woman whom she liked, and her responsibility for influencing the standards of her group was brought to her attention, she became coöperative, responsive and a valuable asset to the group. Though her energies seemed sometimes foolishly directed, there was a definite change in attitude, and in the influence which she exerted. The active-uncontrolled girl is often one of the most useful, socially, and important to the group in which she belongs. Much can be done to control, curb or direct her tendencies, under right conditions.

The "hysterical" girl is more commonly discovered by the room-mate of the offender than by members of the faculty. She constitutes a serious problem, as she readily gains the sympathy of the uninitiated by her temper tantrums, crying spells, "nervousness" and other neurotic symptoms. She must have an audience for her "attacks," so she frequently chooses the time when she actually disturbs the work of a large group of students, who attempt to minister to her needs by giving the desired sympathy. In one group it was quite common for a girl accused of indiscretions and irregular conduct, to retaliate by hysterical crying "fits," in which her violent efforts to secure sympathy served to distract her accusers. In this way she sought to avoid the deserved rebukes of her companions who were entirely justified in demanding that she conform to higher standards of conduct. The reactions of the hysteric can best be dealt with by an experienced person who knows when sympathy should give way to firmness, decision and judgment in dealing with the matter.

The "negatively-suggestible" person is commonly known as "stubborn." Such a girl reacts to suggestion by going directly contrary to the proffered advice, or she makes a negative adjustment in which her parents lose her confidences. Teasing on the part of older brothers and sisters, relatives and even by parents may develop such a latent tendency in any child. It is not that

one "takes after" one or the other parent in this matter! It is much more likely to be an acquired characteristic. This stubbornness is a difficult thing to overcome in adult life, though the knowledge of the danger may assist one in guarding against it.

Tendencies to react in any of the above ways are present in many of us. The reaction may depend upon education, training, environment or immediate stimulus, decidedly more than upon inherent characteristics, social inheritance or ancestry. Certainly no young person should be graduated from our colleges without having these tendencies, when they exist, brought into consciousness, in order that the individual may have help in solving her difficulties, and in overcoming handicaps before they become a part of her habit equipment, personality and fixed attitude towards life.

SUMMARY OF RESULTS

1. From the standpoint of articulation the most frequent causes of failure are the provincialisms and local dialects in New England and the middle-Atlantic groups, the inverted "r" of the middle West and the quality of the Southern vowels.
2. Many minor defects were discovered which indicated personality difficulties, emotional imbalance and ineffective social adjustments (graph II).
3. A high percentage of inaccuracies on s and z sounds and their combinations were found (table 2).
4. On the whole, students from private schools did better than those from public schools and those who had had a little speech training seemed to pass the tests more easily than those who had received none.
5. Negligent lisping or oral inactivity seems more common than organic lisping (this latter being due to some physical defect in the mouth or in the dental condition).
6. Of the nine Freshman students sent home at mid-semesters, for failure in college work, four, or nearly fifty per cent, were classified at entrance to college in the speech correction group.

7. The speech tests give us an opportunity for observing emotional disturbances, feelings of inferiority, peculiarities of personality, nervous instability and so on, which if left uncorrected may interfere seriously with success in college and afterwards. Our chart (graph II) shows how much more common these mental disturbances are than are physical defects in the college girl. The latter have been provided for, for many years by the corrective work of the physical education department. We hope to do an equally effective job with the mental peculiarities.

REFERENCES

(1) BRILL: Quar. Jour. Speech Educ., April, 1923, pp. 129–135.
(2) BRILL: Fundamental Conceptions of Psychoanalysis. Harcourt, Brace Co., New York, 1923.
(3) BLANTON, S., AND M.: Speech Training for Children. Century Co., New York, 1920, p. 77; p. 144
(4) DOWNEY, J. E.: The Will-Temperament and Its Testing. World Book Co., Yonkers, N. Y., 1922, Chap. XV.
(5) PRATT: Mental Hygiene Primer. Publica., 42, Mass. Soc. for Mental Hygiene, Boston, Mass., 1923.
(6) Blanton-Stinchfield Speech Measurements. Pub. by C. H. Stoelting Co., 3037 Carroll Ave., Chicago, Ill.

Additional References

JONES, DANIEL: An Outline of English Phonetics. G. E. Stechert Co., New York, 1922.
JAMES, HENRY: The Question of Our Speech. Houghton-Mifflin Co., Boston.
KLINGHARDT: Ubungen im englishen Tonfall. G. E. Stechert, New York, 1923.
BIRMINGHAM-KRAPP: First Lessons in Speech Improvement. Chas. Scribner's Sons, New York, 1922.
DOLL, E. A., AND LEATHERMAN Z.: Maladjustment Among College Students. Jour. of Appl. Psychol., Dec., 1924.
BURNHAM, W. H.: The Normal Mind. D. Appleton Co., New York, 1924.

WHEN IS A MAN INTOXICATED?

H. L. HOLLINGWORTH

Columbia University

"When is a man baptized?" was once a provocative question. When he has been immersed, or only when he feels a holier impulse, or not until his fellows can testify to his good works? In somewhat the same captious vein the question has more recently been, "When is a man intoxicated?" When he "feels a misery in his breast?" Or only when the neighbors complain of him? Or not until a technical expert discovers a deviation in his reaction time or alcohol in his blood stream? The practical question, as to what does or should constitute the legal criterion, does not concern us here. It is obvious that there are degrees of intoxication, and that the psychologist's notion of it may not accord with that of the policeman.

The question we raise, instead, concerns the way in which various usable criteria are related to the different degrees of toxic influence. Excluding the criteria of the chemist and the physiologist, at least three types of indication are experimentally available in a psychological study of the nature of any drug effect. These may be briefly described as follows:

a. Subjective or introspective report. The individual's verbal communication of the state of his feelings, his array of sensations, impulses, inclinations, and other experiences.

b. General conduct or overt behavior. The individual's attitude, appearance, gait, gesture, facial expression, motor control and the like, capable of report by another observer or witness.

c. Technical measurements. Accurate determination, under controlled conditions, of changes in capacity or variability in the performance of standardised tasks, functions or processes.

In all three ways—introspective account, overt behavior, and technical measurement—may be sought symptoms of the effect of a drug, or, indeed, of any other influence. Of these three methods, which is the most accurate, the most sensitive? Which type of symptom detects the smallest change or betrays the least influence? For one thing, the answer to this question would throw considerable light on conflicting expert testimony in connection with drug effects. Nor is it likely that an *a priori* reply can be anything but a guess. For this reason it seems worth while to report an experimental inquiry into the facts.

Under conditions more fully described in my recent papers on "The Influence of Alcohol," six male subjects were studied for a period of two weeks. During this time they were given, according to a systematic schedule, various dosages. On some days no dose of any kind was given (blank or normal days). On others varying amounts of a control dose were administered (control days). On other days amounts of alcohol ranging from 39 to 78 cc. were given at mid-day, in a brew equal in quantity to the amounts of the control brew.

Each subject kept a "personal data book," in which each day he recorded his general state of well-being, any unusual feelings or sensations, his evening activity, and the amount and quality of his sleep that evening. Five judges (with psychological training) were given these reports, shuffled in random order, each man's report for a given day being on a separate slip of paper. The judges were told that on some days the man had alcohol, on some days the control and on some days nothing. They were asked to sort the slips into these three piles, according to their best judgment of the significance of the reported introspections.

On the other hand, three observers who were present in the laboratory with the men throughout the whole time of their work had every opportunity to note their conduct, conversation, appearance, their attitude toward one another and toward the experimenters, and observed them during the test procedures, but without making comparison of records. At the

end of each day, each observer (two women and one man) reported any unusual signs or symptoms noted in any of the subjects, or else wrote them down as normal or usual in every way. Five judges (those previously described) now arranged slips bearing the separate reports (by all observers for each man for each day) into three piles, according to their judgment as to whether the observer's report indicated alcohol, control, or normal day.

At the same time the objective measurements in an array of tests were secured, and these have been presented in other papers. For the present we are merely interested in the general results of these tests and the question of the degree of drug influence which they were able to detect. In this respect we are now able to compare the three modes of report just described.

By way of illustration of the type of material borne by the slips the following samples are shown, chosen at random. Since there were six subjects and five judges, there were for each experimental day 30 verdicts based on introspection and 30 based on reported conduct. We have distributed the verdicts for each subject, and then combined the results for all six subjects, thus giving the general effect of the day's judgments, based on the introspection or the conduct reports for the day in question. These resemble somewhat the task of a judge or jury, confronted with the defendant's statements on the one hand and the testimony of witnesses on the other hand, the testimony being limited to "observed facts," and excluding inference.

SAMPLE DATA SLIPS—INTROSPECTIVE REPORTS

11-II. Fair this forenoon, sick and nauseated in afternoon. Later in evening very disturbed, dizzy and exhausted. Imagination terrifically active. Wrote a song before going to bed. Slept heavily, usual time.

12-I. Felt lively in forenoon, cheerful and comfortable. About the same in the afternoon, but a little tired. Felt fine in the evening, and took a walk. Sleep good, usual amount.

17-III. Felt good in the morning. Some dizzy in the afternoon, which began to wear away at about 3 p.m. As usual

in the evening, perhaps a little dull. Sleep good, usual amount.

10-IV. Very slight cold, few sneezes during the day. Otherwise as usual. Sleep as usual, good.

17-V. Very active in the morning, with keen desire to make better records. In afternoon slightly exhilarated, and sleepy from 2 to 4 p.m. Feeling was slightly benumbed. Lazy all afternoon. Power of nervous tension seems taken away, without which I cannot make a good tapping record. In the evening, as usual. Sleep, as usual.

SAMPLE DATA SLIPS—CONDUCT REPORTS

11-II. A. Exuberant at 1 p.m., tended to break away from routine (thus brought chair from next room, which he had never done before). Then became sick and mournful, slow, could scarcely steady self for pulse records.

B. Blood shot eyes, can't seem to see right, very slow, holds his head. Better at four o'clock.

C. Frightfully sick, so he couldn't do any tests with a quarter of his usual vim. Complains he never felt more beastly in his life.

11-V. A. Seemed extra cheerful, but no other symptoms.

B. Talked more than usual, inclined to laugh. Unsteady and said he "felt happy."

C. All right, but "silly."

14-III. A. Nothing special noted.

B. As usual.

C. Nothing special noted.

16-VI. A. Said he felt exhilarated but did not show it by talkativeness.

B. Same as usual, perhaps a little less talkative.

C. Acted as usual. No comments.

Table 1 gives the percentile distribution of the judgments for each of the experimental days here studied, when all six subjects are grouped together, and the judgments of the three observers all included. In the first vertical column is given the dosage for a given day. In the other columns are given the distribution of the verdicts, by the two methods, under the three headings.

On the whole the 39 and 52 cc. doses give results closely alike, and the 65 and 78 cc. records are also very similar. But

the results for the former (called here small doses) are very different from those for the latter (called large doses). Comparison of the results can most reliably be based on the four arrays—normal days, control days, small alcohol days and large alcohol days. The following conclusions are suggested by such a comparison of the distributions of the judgments—there being in all 60 judgments for each type of day.

TABLE 1

DOSAGE	BASED ON INTROSPECTIVE REPORTS			BASED ON BEHAVIOR REPORTS		
	Alcohol	Control	Normal	Alcohol	Control	Normal
	per cent	*per cent*	*per cent*	*per cent*	*per cent*	*per cent*
Normal record, two days.....	13	3	84	7	0	93
Control record, two days.....	10	57	33	5	10	85
Alcohol doses, small:						
39 cc.....	60	30	10	23	10	67
52 cc.....	50	23	27	33	47	20
Average.....	55	27	18	28	28	44
Alcohol doses, large:						
65 cc.....	83	7	10	77	13	10
78 cc.....	70	10	20	73	17	10
Average.....	77	8	15	75	15	10

Blank Days are correctly judged as "normal" with less than 10 per cent error, the errors being very slightly biased toward the alcohol verdict. Judgments are about equally reliable, whether based on the introspective accounts or on the conduct reports.

In the case of the control days, judgments based on the conduct reports are not distinguishable from the verdicts rendered on blank days. But in the case of judgments based on introspective accounts, the control days are diagnosed correctly in considerably more than half the number of cases. Based on

these accounts the control days are just about as correctly identified as are the "small alcohol" days, on the same basis.

With "small alcohol" doses, about half the judgments based on the introspective accounts will be for and half against the verdict "alcohol." Of the half against alcohol, about equal numbers favor control and normal. That is to say, three-fourths of the judgments agree that "something is wrong," but only half of the total judgments correctly identify the condition. Based on the conduct records, however, the verdict is against alcohol, in a ratio of about 3 to 1. The non-alcohol votes here favor normal, but the votes for normal in this case are only half as frequent as on actual normal days. Judgments based on subjective account are thus more reliable than those based on behavior record, for the small doses. That is to say, alcohol affects the feelings before it does the grossly observable conduct.

With "large alcohol" doses, the votes are overwhelmingly for alcohol, whether based on subjective account or on behavior report. And the distribution of the judgments is quite alike in the two cases.

Generalizing, we may say that, if a majority vote be allowed to decide for or against intoxication, the small alcohol doses do not produce it when the judgment is based on the behavior reports. When the verdict is based on the introspective account, the alcohol effect is just barely said to be present.

Much the same thing is to be said for the control days, except that here the behavior reports are of no value at all and control days are confused with normal rather than with alcohol days. Large alcohol days are correctly judged by a three-fourths vote, whether based on introspective account or behavior report.

During the course of the experiment, measures were made of ability in steadiness, tapping, coördination, color naming, adding, opposites, substitution, paired associates, and pulse rate. Every one of these measured traits or processes reveals the effect of the "small alcohol" doses, positively and unambiguously. This effect the behavior reports failed to

reveal, and the introspective accounts disclosed it only equivocally and by a narrow margin.

In fact, in a special experiment, covering one day, in which the effect of added dosage was studied, beginning with 13 cc., four of the technical tests, viz., adding, substitution, opposites and paired associates, began to show definite alcohol effects after only 26 cc. These are the most clearly verbal, abstract, symbolic or mental of the processes examined.

It seems clear then that effects may be present which are measureable by properly chosen technique long before they are manifested in the individual's gross conduct in a normal, social, working situation, and even a considerable time before they are indicated by the subjectively reported symptoms.

The order, in decreasing accuracy, is as follows:

A. Testimony of technical measurements
B. Evidence of subjective or introspective accounts
C. Judgments based on general conduct or overt behavior

In the present experiment the third of these identifies only the large dose effects, the second correctly identifies the small dose effects, whereas the first correctly identifies effects from doses smaller than any of these here considered. In the case of alcohol, at least, the effects will need to be of that order of magnitude produced by what are here called "large doses" (65 to 78 cc.), consumed within a period of not more than one hour, before all three sorts of testimony agree.

It is at least clear, at any rate so far as alcohol is concerned, that the technique of experimental psychology is capable of detecting drug effects considerably smaller than those revealed by introspective testimony or by general observations of appearance and behavior. Whether or not "toxic effects" are present can be determined only after the method of their detection is specified. Until such specification is agreed upon or conventionalised, we may expect expert witnesses to disagree and judicial decisions to be uncertain.

Finally, it is quite possible that the value of these three modes of detection will vary with the drug concerned. That

drug effects, of any kind, will be shown in gross conduct more readily than in the technical measurements, seems quite unlikely. Whether or not some drug effects show themselves more reliably in introspective account than in either of the two other methods, is at least problematic. It is also entirely possible that the method best suited for one individual is not the most sensitive for another—that autogenic reinforcement, for example, might enable one individual to maintain his technical performance and also to control his appearance and general conduct, although a frank introspective account would reveal the presence of a disturbing influence. In the present instance we have ignored such individual differences in favor of an indication of the general tendency in the case of the three modes of detection.

Definitely supporting these conclusions concerning the relative accuracy of the three methods, are the results of some experiments which I reported about twelve years ago. In this case a study was made of the accuracy of performer and of witness in judging whether a test performance just completed was "usual," "worse than usual" or "better than usual." Since actual measurements were also made, but kept from the knowledge of the judges, we have in this instance also the three techniques—performer's introspection, observation of witness, and technical measurement.

Considering all those cases in which the judgments were rendered with anything but the lowest degree of confidence (A and B judgments) the per cent correctness of the verdicts by the three methods were as follows:

A. Testimony of technical measurements.	100 per cent correct
B. Introspection of performer.....	84 per cent correct
C. Judgment of witness.....	80 per cent correct

The order of relative accuracy is quite the same as in the case of the present drug experiments.

It is not my purpose to impugn the practical value of the usual testimony. It may well be that for legislative purposes the most useful criterion is to be found in the rough observa-

tions of a witness, unsupported either by introspective accounts or by technical measurements. But this method we have shown to be the least accurate and the least sensitive of the three techniques here studied. When it is concluded that "there was not the least sign of intoxication;" that the drink "did not in the slightest degree impair the mental faculties or bodily functions;" and that "none manifested the slightest symptom," it is at least scientifically, and perhaps practically important to realize that the presence or absence of a symptom is obviously a function of the technique employed in searching for it. In the present connection we have been interested merely in the determination of the relative accuracy of three such techniques.

THE CLINICAL VIEWPOINT IN VOCATIONAL SELECTION[1]

MORRIS S. VITELES

University of Pennsylvania

This paper grows, in part, out of a conference with the works manager of a metal working plant. This gentleman was faced with a production problem in three departments: (1) a press room, (2) a milling and drilling room, and (3) an assembly department. He had developed an almost naïve faith in the potency of psychological tests for solving this problem by selecting workers best qualified for each of these three types of work. In discussing the proposed tests the works manager insisted upon placing the following limitations upon the work of the psychologist.

1. That the tests be limited to between thirty and forty-five minutes.

2. That only machines and material found in the plant be used in the manufacture of the apparatus involved.

3. That the tests, when devised, be capable of administration and interpretation by a clerk in the employment office.

It is the third limitation which is to be discussed in this paper. The first impulse is to criticize the works manager for insisting upon this limitation. But it requires only a hasty glance over the literature of vocational selection to discover that he was merely insisting upon a condition which the vocational psychologists have themselves established. Again and again, both in technical articles and in advertising announcements, have these psychologists repeated that these tests can be given, graded, and evaluated by a minor clerk in the employment office. Only

[1] A paper read at the meeting of the Industrial Section, American Psychological Association, Washintgon, December 30, 1924.

rarely has it been suggested that the objective scores of vocational tests are at best uncertain diagnostic criteria; that these must be supplemented by an observation and analysis of performance by a trained examiner.

For example, in a lengthy exposition on "Measurement in Vocational Selection: an Outline of Research Procedure," Freyd states that[2]

> it requires more training to conduct an experiment in vocational tests than to administer the tests. An economy of effort will result if the experimenter assigns the task of administering the tests to another. He should insist on the selection of a tester who is fairly intelligent, polite and tactful, and favorably inclined toward the theory of tests and scientific method in general.

In a later section of the same article he says that "the maintenance of test procedures and other measurements in vocational selection clearly demand an oversight of a carefully statistically trained and judicious-minded person." It is important to note that in neither the case of the supervisor of test procedures nor of the administrator of tests is training in psychology mentioned or even implied. Apparently, politeness and statistics are more important for successful work in vocational selection than psychology. Moreover, it is the duty of the overseer to supervise the maintenance of test procedure. In no place is it suggested (and Freyd's viewpoint in this connection is a representative viewpoint) that it may also be the function of psychologists working in industry to evaluate and interpret not only the score, but the actual test performance of the individual applicant for employment.

The insistence upon this point of view is based upon a number of assumptions.

1. The first is a statistical assumption. It says that[3]

[2] Max Freyd, Measurement in Vocational Selection, Journal of Personnel Research, vol. II, nos. 8 and 9, December, 1923—January, 1924, p. 377.

[3] Max Freyd, Measurement in Vocational Selection, Journal of Personnel Research, vol. II, nos. 8 and 9, December, 1923—January, 1924, p. 272.

the test score in the practical situation is always the known variable and from it the probable degree of vocational success may be predicted on the basis of the known relationship of test scores and degrees of vocational success. According to this assumption the evaluation of the tests consists in the comparison of two series of measurements made of the subjects of the experiment, the test scores and the criterion of success.

There is no time here to criticize in detail this assumption. Those of us who are interested in the examination of children for educational purposes, for correctional purposes, or for purposes of guidance have found it impossible to diagnose and treat on the basis of objective score alone. It has been found necessary to supplement the objective data of mental measurement with the observation and analysis of performance. Mental ages, I.Q.'s, and time scores have been found to be useful in guiding the examiner, but not in directly indicating the intellectual level or specific abilities of children.

There is a qualitative aspect of mental analysis which goes hand in hand with quantitative analysis. Maxfield points out that[4] "in the use of standardized mental tests subjective judgments of the experienced examiner in regard to the qualitative aspect of the subject's reactions, whether verbal responses or other types of behavior, are significant."

The clinical psychologist has been forced by experience to place himself in the position of a physician in the matter of diagnosis and recommendation. With reference to diagnosis the psychological test may, in a sense, be compared to the various objective tests, some of them laboratory analyses, made by the physician's assistant. The physician permits his assistant to test the temperature of the patient, to read his pulse; he seeks the laboratory technician's report on the blood count, the incidence of sugar in the blood, the rate of oxidation, etc., but in practically no instance does he throw upon the assistant techni-

[4] F. N. Maxfield, The Use and Abuse of Standard Intelligence Tests in Individual Examinations. Reprinted from the Proceedings of the 48th Annual Session of the American Association for the Study of the Feebleminded, 1924, p. 21.

cian the responsibility of a diagnosis. Diagnosis involves an interpretation of the objective data—interpretation involving scientific knowledge which the assistant does not possess, and which, therefore, makes him incompetent to render a diagnosis. There are very few exceptions; the technician's report of a four plus Wassermann establishes without doubt the presence of syphilis, and in a sense dictates the diagnosis and the treatment, but even in medicine these specifics are rare, and in the field of psychology the test which is a specific diagnostic measure is still rarer, if it exists at all.

"Many psychologists," as Maxfield[5] wisely observes, "have had the tendency to confuse mental measurement with analysis and mental interpretation." In vocational selection in industry this has especially been the case. Statistics and the statistical point of view have dominated. It is the opinion of the writer that in the cause of greater scientific accuracy in vocational selection in industry the statistical point of view must be supplemented by a clinical point of view. It must be recognized that the competency of the applicant for a great many jobs in industry, perhaps even for a majority of them, cannot be observed from an objective score any more than the ability of a child to profit from one or another kind of educational treatment can be observed from such a score. There is no reason for suspecting that the capacity of an individual motorman to avoid accidents, or of a printer's apprentice to profit from instruction in this trade can be expressed in an objective score, as easily interpreted by a minor clerk as by a trained psychologist, than for suspecting that the mental status of a child is revealed in the I.Q. which can be obtained by any teacher who owns a copy of Terman's Condensed Guide and a set of testing material. The one problem is as complicated as the other; the objective score in one case has in it as many elements of error as in the other, and an adequate diagnosis in both involves interpretation by a trained psychologist based on observation of performance and a consideration of related data.

[5] Id., p. 23.

2. Another assumption which has led to the domination of the statistical point of view in vocational selection is that industry is not ready to pay the cost of a more refined and more detailed method. This assumption is perhaps justified in fact. The strings of industry's purse are sensitive, and it is probably true that the envisagement of a too ambitious program of vocational selection, even if it were endowed with more of the quality of scientific accuracy, might well result in the absolute closure of the purse, and the stoppage of the slight dribble of coins which have been carefully counted out for psychological work in industry. This is a real danger, but it hardly justifies the easier and the less scientific course. Education of industrial managers to the problems and needs of adequate research and the encouragement of endowed research are the ways of meeting this contingency. There may be a delay in the development of the application of psychology in industry, a diminution in the quantity of work, but this is preferable to the defense of a position which will prove ultimately to be untenable.

Industry should be told quite frankly that there are some jobs for which workers can be satisfactorily selected by means of tests graded and statistically evaluated by minor clerks in the employment office. The specific jobs for which this holds true are determined in part by the duties of the job, the universality and what might be termed the superficiality of the qualities needed for success in the job, and by certain economic factors such as cost of labor turn-over, the relative turn-over of the total group employed in this type of work, the time and cost of training, etc. But industry must also be told with equal frankness that there are a great many jobs for which workers can not be adequately selected by tests administered and interpreted by employment clerks. It must be told that the selection of applicants for such jobs involves an examination by a trained psychologist who depends upon his scientific knowledge of human behavior as well as upon the test results. Industry must be told that the immediate cost of such a procedure may be great, but that ultimate economy is implied in the acceptance of such a procedure.

3. A third assumption, implied rather than stated, an assumption which is perhaps not recognized by a great many workers in this field, is that of the relative unimportance of the individual worker in industry. A selective process which is satisfied with the probable adequate selection of a group of workers, rather than with the proper placement of the individual worker (and such is the case n selection based upon objective scores alone) fails to give adequate consideration to the well-being and interest of the individual worker, an interest which in comparison with the point of view of European psychologists, has been shamefully neglected by the American psychologist.

Other assumptions can be stated, but they can not be presented in a paper of this length.

A discussion such as this of the clinical viewpoint in industrial selection involves a reconsideration of the place of the psychologist in industry. The acceptance of the statistical viewpoint makes place in industry for what has been called a psychotechnician, a worker with a minimum training in psychology and a maximum training in statistics. He is a technician in the fullest sense of the word, well trained in method, but not in theory. He understands the various steps to be taken in standardizing a test and in statistically treating the results for the revelation of significant scores. In many cases he need know little about mental tests, since statistical treatment will reveal the significant ones, although a little information about tests may be helpful in providing shortcuts in the determination of the significant ones. Of clinical observation and interpretation of performance he need know nothing at all, and generally knows nothing. He is a consultant in industry, supervising the construction and statistical evaluation of tests and the administration of such tests by a polite, but not necessarily intelligent, clerk.

If the clinical point of view is to affect vocational selection the worker must be more than a psycho-technician. He must be a psychologist trained in psychological theory and methods as well as in statistics. He must be more than a consultant; he must be a co-worker in the personnel department, imbued

through experience and study with an understanding of the selection problems of the particular plant and with the problems of the job for which he is testing. He must function in the examination of the applicant, applying his experience in the observation and analysis of performance in deciding whether a given applicant should be employed. He becomes in a sense a sort of superior employment officer, weighing the test score, his observation and analysis of performance with other essential data about the applicant in determining whether or not the applicant is competent for the job. His judgment is a diagnosis, as that of a physician, based upon a consideration of all the data affecting success or failure on the job.

These are the two extreme positions. There may be all sorts of compromises, but the necessity of the active functioning of the psychologist in the employment of the individual worker is a necessary concomitant to the adoption of the clinical viewpoint in vocational selection.

This does not mean that there remains in industry no place for what has been called the psycho-technician in industry. He has exactly the same place in industry as the laboratory technician has in medicine. He is the assistant to a trained examiner, providing such objective quantitative data as he can for interpretation by the trained examiner. In certain cases, as has been indicated, he may even be the sole representative of psychology in industry, functioning in the standardization of tests and in the determination of a set of statistical values useful in employing a particular test in selection. Such a job a psycho-technician can very well do, but his services should be provided with the understanding that industry is receiving not psychological services, but technical services; that it is using in employment not necessarily psychology but statistics slightly diluted with a smattering of psychology. Such services should not be sold under false pretenses, as has often been the case.

This paper is limited to the selection of workers for a given job. The neglect of the clinical viewpoint is equally apparent in the use of tests in vocational guidance. In this field, where a correct reading of the mental abilities and defects of the

individual is perhaps more important than in industry, the need for the clinical viewpoint becomes even more urgent. It is perhaps impossible to overestimate, both from the point of view of the individual studied and the progress of applied psychology, the harmful effect, for example, of such a logical development of the statistical point of view in guidance as the program for long-distance guidance by correspondence, an advertising prospectus of which has recently been widely distributed by a well known publisher of psychological tests. It is impossible to discuss this aspect of the application of the clinical viewpoint in vocational selection within the limits of this paper, but the writer takes the opportunity to suggest the need for a re-evaluation of testing projects in vocational guidance in the light of the point of view presented herein.

ANALYZING PERSONALITY

PART I

GEORGE C. BRANDENBURG

Purdue University

How often one hears it said that Mr. A or Mr. B is a very bright man, but that he has no personality. Sometimes it is even said that such men as scientists and intellectual persons as a class are characterized to a certain extent by poor personalities. And occasionally one hears the converse, that Mr. So-and-so has a fine personality, but that there is really little to him, that he lacks depth. Promoters, reformers, and salesmen are not infrequently so described. The prominent rôle which this term personality plays in our every-day affairs is further indicated by the fact that in practically all recommendation blanks there are numerous questions concerning the applicant's personal qualities. Does he have a good personality? Is he tactful in dealing with other people? Is he forceful? Has he a positive manner? Does he coöperate readily with others? Is he industrious? Scores of such terms as occur in these questions are in common use in the literature of employment agencies and are indicative of the interest which employers generally have in other phases of the individual's make-up aside from his immediate ability to do the work in question.

What does the term personality really mean? What elements enter into it and how are they related to one another? How is personality related to intelligence? What does it have to do with success in various fields of activity? These are some of the questions which must have occurred to the reader time and again and they are likewise questions which must frequently occur to any thoughtful student of human affairs.

We use the term as if its meaning were so well understood that it required no definition, but such is far from the case. Furthermore, we frequently in our discussions use such terms as personal and character traits as if they were synonymous and could be used interchangeably. Again this is not quite true. If we examine psychological literature we find that the term, personality, has a wide variety of meanings, depending on the viewpoint of the writer. According to Warren it is "the entire mental organization of a human being at any stage of his development." Woodworth regards personality as including one's physical, temperamental, instinctive, and intellectual qualities. Watson, thinking of personality entirely from the standpoint of a behaviorist, says, "Let us mean by the term personality an individual's total assets (actual and potential) on the reaction side. By assets we mean, first, the total mass of organized habits; the socialized and regulated instincts; the socialized and tempered emotions; and the combinations and interrelations among these; and second, high coefficients both of plasticity (capability of new habit formation or altering of old) and of retention (readiness of implanted habits to function after disuse)." A most interesting and extreme view of the meaning of personality is held by those few writers who believe that it is explained wholly in terms of the activities of the ductless or endocrine glands. Of course, these investigators are more interested in showing that the glands are the basis of personality and especially of its variations, than in any definition of the term.

As used by the writer the term is understood to indicate a composite of an individual's typical reactions, physical, intellectual, emotional, to his environment, together with his various physical characteristics which constitute what we call his general appearance. Some writers go still further in this direction and assert that the term represents the sum total of the individual's reactions during his existence as a unified organism. It seems quite useless, however, and impractical to make the term so all inclusive as this; for it is quite obvious that in our actual perception of the personality of another we are limited to a frac-

tional part of his total behavior. In our ordinary judgments of the personality of an individual we make use of only a few cross sections of his behavior and physical characteristics. Naturally we realize that such judgments are inadequate, but in any event they are what we use in practical every day affairs and must no doubt continue to use.

A little reflection, then, concerning the meaning of one's personality makes it clear that it depends upon both his inherited qualities and the environmental conditions which have surrounded his existence. We shall not be particularly concerned with the relative influences of these two factors in the development of personality. It should be noted, however, that one's environment has somewhat more to do with his personality than with his character; while heredity apparently plays the more important part in the matter of character.

PERSONALITY AND CHARACTER

The term character is generally used to indicate a somewhat restricted part of a person's individuality and commonly refers to those emotional and instinctive qualities which are largely innate. In any case it has little reference to the physical characteristics or the appearance. Such qualities as honesty, reliability, sincerity, and moral uprightness come to one's mind in connection with the term character. In general, then, the term personality is somewhat more objective, more akin to the term reputation and has reference to the manner in which the individual's behavior and appearance appeal to his friends and associates; it is more dependent on social judgment while character is more subjective and less susceptible to social evaluation. The term personality is associated with such qualities as enthusiasm, forcefulness, address, tact, and physical appearance in general. But there is much overlapping in the terms character and personality so that it is quite impossible to make any clear-cut distinction between them. When a sufficient number of critical studies have been made to form an adequate basis, it will, no doubt, be found desirable to set up fairly specific limitations of these terms. In any study of personality

TABLE 1

Showing the intercorrelations of twenty-three personal traits

	ACCURACY	ENTHUSIASM	AGGRESSIVENESS	SELF-RELIANCE	MEMORY	POPULARITY	MOTOR ABILITY	TACT	GENERAL ABILITY	RELIABILITY	COÖPERATION	REASONING ABILITY	GENERAL INFORMATION	ORIGINALITY	SYMPATHY	SPEED IN WORK	SOCIAL AND CIVIC INTERESTS	ADDRESS	SINCERITY	INDUSTRY	NEATNESS	APPRECIATION OF HUMOR	MORAL HABITS	TOTAL PERSONALITY
Accuracy		72	86	72	83	68	67	62	70	66	72	71	58	76	64	78	72	48	33	60	33	35	36	*87*
Enthusiasm	72		78	71	72	76	64	81	80	66	67	65	85	66	58	71	61	50	38	53	35	28	10	*86*
Aggressiveness	86	78		61	72	73	67	69	53	69	52	60	53	69	57	78	62	49	25	48	45	48	30	*85*
Self-reliance	72	71	61		70	59	71	66	80	83	76	80	66	52	45	73	34	20	77	39	18	−02	12	*85*
Memory	83	72	72	70		58	64	62	47	73	52	61	50	69	66	66	47	34	54	69	27	35	41	*85*
Popularity	68	76	73	59	58		75	90	76	72	80	61	70	61	84	44	67	61	43	26	56	46	02	*84*
Motor ability	67	64	67	71	64	75		62	76	53	60	76	54	58	58	56	40	52	61	26	22	26	16	*81*
Tact	62	81	69	66	62	90	62		79	64	80	56	72	52	67	61	57	53	36	25	54	44	13	*80*
General ability	70	80	53	80	47	76	76	79		72	83	85	68	56	64	61	44	49	50	29	41	25	−24	*80*
Reliability	66	66	69	83	73	72	53	64	72		71	70	56	47	62	68	46	27	81	25	15	12	19	*79*
Coöperation	72	67	52	76	52	80	60	80	83	71		73	64	60	67	40	35	33	53	18	33	18	−27	*78*
Reasoning ability	71	65	60	80	61	61	76	56	85	70	73		69	57	41	66	32	47	62	26	14	−15	−18	*77*
General information	58	85	53	66	50	70	54	72	68	56	64	69		54	48	72	48	41	47	68	32	−08	20	*76*
Originality	76	66	69	52	69	61	58	52	56	47	60	57	54		71	53	70	72	45	54	12	30	13	*75*
Sympathy	64	58	57	45	66	84	58	67	64	62	67	41	48	71		28	68	54	36	33	19	52	13	*73*

Speed in work	78	71	78	72	66	44	56	61	61	68	40	66	72	53	28		48	26	52	47	32	05	26	*71*
Social and civic interests	72	61	62	34	47	67	40	57	44	46	35	32	48	70	68	48		69	−03	29	45	47	20	*61*
Address	48	50	49	20	34	61	52	53	49	27	33	47	41	72	54	26	69		−17	13	17	45	−04	*55*
Sincerity	33	38	25	77	54	43	61	36	50	81	53	62	47	45	36	52	−03	−17		26	12	−20	33	*54*
Industry	60	53	48	39	69	26	26	25	29	25	18	26	68	54	33	47	29	13	26		−17	12	60	*52*
Neatness	33	35	45	18	27	56	22	54	41	15	33	14	32	12	19	32	45	17	12	−17		−02	−22	*43*
Appreciation of humor	35	28	48	−02	35	46	26	44	25	12	18	−15	−08	30	52	05	47	45	−20	12	−02		36	*32*
Moral habits	36	10	30	12	41	02	16	13	−24	19	−27	−18	20	13	13	26	20	−04	33	60	−22	36		*17*

then, it is obvious that we shall be dealing with such traits as enthusiasm, aggressiveness, and neatness in appearance which are clearly qualities of personality. But since personality includes character, it will also be found desirable to include such outstanding qualities of character as reliability, sincerity, and moral character. At present we know comparatively little concerning the tendencies of such personal and character traits as those just mentioned to be mutually attractive, antagonistic, or neutral in one's make-up.

What likelihood is there, for example, that the intelligent person will be good, careful, tactful, coöperative, or reliable? What are the chances that the enthusiastic, and optimistic man will be sincere, sympathetic, or broadminded? How far may we expect to find the individual with an outstanding personality possessing such traits as common sense, self-reliance, and originality? To what extent is a student's grade determined by his personality? How far is a man's income dependent upon his personality?

AN EXPERIMENTAL STUDY

Group I

In an endeavor to throw some light on such questions as those just suggested, twenty-nine senior engineering students in a vocational psychology class rated each other with respect to twenty-three personal and character traits. At the close of the rating process and without any knowledge of its results, the class was given the Army Alpha Intelligence test. The traits used in the rating experiment were selected from a list of some one hundred fifty traits assembled from various rating and personnel systems. The selection was made through a process of preferential voting by the students so that those selected were judged to be the most important ones for a good personality. These traits are: Accuracy in work, enthusiasm, aggressiveness, self-reliance, memory, popularity, motor ability, tact, general ability, reliability, coöperativeness, reasoning ability, general information, originality, sympathy, speed in

work, social and civic interest, address, sincerity, industry, neatness in appearance, appreciation of humor, and moral habits. Some of these traits, for example, general information, can hardly be called personal traits since they represent acquisitions of knowledge and the like rather than fixed tendencies or qualities. No attempt has been made to avoid this apparent inconsistency or to reduce the list to simple, clearly definable qualities. This is probably a weakness in the experiment. The chief justification of the traits included is that they are frequently used in rating and personnel work, are well understood, and are therefore assumed to be significant.

After the traits had been selected each member of the class agreed to rate each of the other members by assigning him a rank in each trait which would represent his standing in that particular trait with reference to the other members of the group. The rating was all done in the classroom and consumed on the average about two and one-half hours time. There was no conferring among students after the rating began and all the results were held by the instructor as confidential. After the rating sheets had been handed to the instructor the estimates were all summarized so that each student received a final ranking in the group in each trait. This final rank thus represents the pooled judgments of twenty-eight of his associates concerning his relative merit in each of the qualities. It was then possible to assign each student a final rank not only in each trait but in the total of all the traits or in personality as a whole. For the purposes of the experiment we assume that this is a correct and true measure of each individual's personality with reference to this particular group.[1]

[1] An individual's rank in any trait is thus the result of the combined judgments of twenty-eight of his classmates and since he was rated in twenty-three different traits, his total personality rank, that is his final rank, is based on a combined total of 644 separate judgments. It should be recalled also that twenty-seven of these twenty-nine men entered the Engineering School of Purdue University in September, 1915, and graduated in June, 1919. Two of the men entered in September, 1916, and finished in June, 1919. Thus, with the two exceptions, the men were in school together for four years. Just how intimately

As has been pointed out elsewhere there is no way of determining how far these estimates really represent valid measures of the various personal qualities. We can determine, however, by approved statistical procedure how reliable or consistent they are; that is, how far the same ratings would be again obtained under the same conditions. We have computed the coefficients of reliability for ratings made in other groups using a somewhat different array of traits.[2] We shall refer to the details of this matter a little later. It may be noted in passing that the reliability, while varying somewhat for different traits, is generally found to be high where the number of persons rating a given trait is as large as it is in this case.

Table 1 shows the intercorrelation among the twenty-three traits as well as the correlation between each and the total personality as indicated by the final rank of the individual in the various qualities. The traits are arranged in the table in the order of size of the coefficient of correlation with total personality. Referring to the table we may then conclude that the most important of these twenty-three traits as factors in the make-up of the individual's personality is accuracy in work, with a correlation of 0.87; the second most important is enthusiasm 0.86; while the least important so far as personality is concerned is moral habits, with the negligible correlation of 0.17.[3] On first thought it may seem somewhat surprising

the members of any group of students know one another cannot well be determined. There is much reason to believe that the men who coöperated in this rating experiment were sufficiently well acquainted to give their pooled judgments of one another a considerable degree of validity. Of course, their judgments of any individual do not agree; such judgments never do, and no degree of familiarity would produce complete agreement. In nearly every case an individual is found to be rated in a given trait all the way from the very best to the very poorest, thus indicating the unreliability of a single individual's judgment of another. This, however, is in itself no proof that the combined judgments are not highly valid.

[2] The reason the reliability was not calculated for the ratings in this particular group was that as soon as the estimates had been summarized, the original rating sheets were destroyed.

[3] In making the assumption that a given trait is relatively impor-

that moral habits as a quality is shown to be practically irrelevant in the matter of personality so far as students are concerned. Two reasons for this have occurred to the author: first, there is the characteristic reaction of the average student to the display of moral habits by another student. In spite of the instructor's warning to the contrary these students were probably inclined to think of moral habits as regular attendance at Sunday School and church, abstinence from the use of tobacco, non-use of profanity, etc. Naturally we should expect little correlation between moral habits, in such a sense, and excellence of personality. Second, it is quite possible that moral habits in the best and broadest sense of the term is not so much of a factor in a good personality as it is sometimes supposed to be. In any case it must be remembered that it is a positive factor even though it has little weight; in other words, moral habits do contribute slightly to, rather than detract from a good personality. In later experiments it has been found that the substitution of the term moral character defined as uprightness and honorableness, for moral habits increases the correlation

tant in the constitution of personality according to its index of correlation, I am fully aware of the existence of an error which runs throughout any such series of inter-correlations. This error which is always present in the correlation of a series of measures or ratings with a second series in which the first is included, naturally tends to raise all the correlations somewhat. However, with the fairly large number of traits used in this experiment, the influence of any given trait on the composite is not likely to be sufficient to affect the correlations very materially. In any case it is unlikely that the relative rankings of the various traits vary much from what they would be if the two series correlated were entirely discrete. Furthermore, since the purpose is to discover suggestive trends in the personality structure rather than exact relationships, it is obvious that the presence of this error does not justify serious criticism of the method employed.

The failure to introduce further refinement of statistical procedure is also explained by the fact that we are interested in the general nature of personality. It is quite likely that the employment of refined and complicated statistical procedure in dealing with measurements of the kind with which we are here concerned would tend to confuse rather than to clarify the situation.

with personality in a considerable degree. One is rather disappointed also that such a desirable quality as industry should not receive a more cordial welcome in this group of attractive personal qualities. We may infer that while industriousness is

TABLE 2

Agreeable traits

	NUMBER OF VOTES
1. Honesty, fairness in grading	59
2. Courtesy and consideration	53
3. Liberal and progressive attitude	47
4. Cheerfulness	46
5. Clearness	35
6. Punctuality	18
7. Generosity	17
8. Cleanliness and neatness	15
9. Morality	7
10. Intellectuality	6

TABLE 3

Disagreeable traits

	NUMBER OF VOTES
1. Lack of courtesy in the classroom	88
2. Unfair practices in grading, etc	69
3. Disagreeable qualities of disposition	42
4. Egotism	30
5. Irritating personal habits and peculiarities	29
6. Lack of cordiality	26
7. Lack of punctuality	20
8. Lack of dependability	18
9. Lack of public spirit	15
10. Untidiness and carelessness in dress	13

a very valuable asset from some standpoints, it does not tend greatly to give one an outstanding personality.

An examination of the table reveals some interesting tendencies in the interrelationships of the twenty-three traits of the individual's make-up. Perhaps the outstanding fact is the

tendency of all except the last three traits to correlate to a fair or a considerable degree with each other as well as with total personality. This bears out generally the results of other studies in this respect and warrants the conclusion that desirable traits of character tend to accompany one another.

Assuming, then, as we have for the purposes of our study that these twenty-three traits make up one's total personality and that the combined estimates are reliable, we may say that the excellence of an individual's personality is dependent on the various traits roughly in the degree indicated by the following classification:

Largely by:	Somewhat by:	Slightly by:
Accuracy in work	Reliability	Neatness
Enthusiasm	Coöperation	Appreciation of humor
Aggressiveness	Reasoning ability	Moral habits
Self-reliance	General information	
Memory	Originality	
Popularity	Sympathy	
Motor ability	Speed in work	
Tact	Social and civic interest	
General ability	Address	
	Sincerity	
	Industry	

In a similar manner the relationship of any given trait to each of the others may be shown. Popularity, for example, is determined by other qualities in this group thus:

Largely by:	Somewhat by:	Slightly by:	Not at all by:
Tact	General ability	Appreciation of humor	Moral habits
Sympathy	Enthusiasm	Speed in work	
Coöperation	Motor ability	Sincerity	
	Aggressiveness	Industry	
	Reliability		
	Accuracy in work		
	Social and civic interest		
	Address		
	Originality		
	Reasoning ability		
	Self-reliance		
	Memory		
	Neatness		

Of course, such classifications of relationships are only suggestive: and it must also constantly be borne in mind that we are

here dealing with a very select group; that is, these students had all practically completed four years of training in the same engineering school, were all graduates of public high schools, and had all come from somewhat similar home environments. Whether the trends in relationships between personal traits which we have found to exist in such a group would be found in a more representative population containing all grades and types of individuals can only be determined by further studies.

As a further check on the question of how various traits are related to popularity, each of forty seniors was asked to submit a list of ten qualities which he found most attractive and agreeable in his instructors; a few days later he was requested to submit another list of ten qualities which he found most disagreeable in instructors. A total of one hundred five agreeable and ninety disagreeable traits appeared in their submitted lists. Related traits were combined and the lists then summarized under ten headings for the agreeable traits and ten for the disagreeable. Tables 2 and 3 in which the various classes are arranged in order of their importance according to the number of times they were mentioned show the results.

These tables show the same general arrangement of qualities with reference to the popularity of an instructor as does the previous list with relation to popularity in general. It is apparent from both studies that such qualities as intellectuality, moral habits, industry (not mentioned in the latter lists), and neatness have little to do with one's popularity. The old adage "Virtue is its own reward" seems entirely applicable in this case. Consideration for and sympathy with others, courtesy, a coöperative spirit, enthusiasm, and cheerfulness, fairness and reliability are the traits which are most likely to make one a favorite with his associates. One would suppose that intellectuality in an instructor would surely be one of the first qualities a student would look for and admire; but after all students are only human and apparently crave first the same humane qualities in their instructor that they demand in their classmates.

Group II

One year after the study just described a second group of twenty-five senior engineering students rated each other following the same technique as group I, with the exception that nine traits were used instead of twenty-three. The traits used in the second group are by definition made somewhat more inclusive than were those formerly used. They are: common sense, judgment, or tact; trustworthiness, reliability, or dependability; executive ability, planning and directing; general

TABLE 4

	COMMON SENSE	ORIGINALITY	AGGRESSIVENESS	EXECUTIVE ABILITY	BROAD-MINDEDNESS	SELF-RELIANCE	INTELLIGENCE	RELIABILITY	MORAL CHARACTER	TOTAL
Common sense		93	88	90	83	84	88	84	65	*95*
Originality	93		94	92	83	92	87	72	52	*95*
Aggressiveness	88	94		89	83	92	86	74	56	*94*
Executive ability	90	92	89		79	88	85	85	51	*92*
Broad-mindedness	83	83	83	79		83	73	83	57	*90*
Self-reliance	84	92	92	88	83		88	70	49	*89*
Intelligence	88	87	86	85	73	88		70	50	*88*
Reliability	84	72	74	85	83	70	70		74	*85*
Moral character	65	52	56	51	57	49	50	74		*66*

intelligence, mental caliber, capacity to learn; resourcefulness, originality, or initiative; aggressiveness, energy, or industry; broadmindedness, tolerance, open-mindedness; confidence or self-reliance; moral character, uprightness, honorableness.

Table 4 shows the intercorrelation among these traits together with the correlation of each with total personality. It will be noted that the correlations are uniformly somewhat higher than those in the first group, but that the general trend is similar. Those qualities which in their display compel attention, the qualities of enthusiasm, aggressiveness, and originality when combined with common sense, tact, and accuracy in work,

are most prominent; while moral character is of lesser prominence in the outstanding personality. The fact that moral character shows a fairly good correlation with total personality in this group whereas moral habits showed a negligible one in the first group is probably an indication of the difference between the interpretation of the two terms by the raters. Moral character is presumably a more acceptable term as a personal quality to the student than is the term moral habits.

While it is entirely likely that there is a considerable tendency for such desirable traits as we are here considering to accompany each other in one's personality, it is very doubtful that they are

TABLE 5

Showing the reliability coefficients of the ratings in nine personal qualities in a group of twenty-five students

QUALITIES	RELIABILITY COEFFICIENTS
General intelligence	0.87
Aggressiveness	0.85
Executive ability	0.75
Originality	0.74
Reliability	0.73
Common sense	0.67
Moral character	0.64
Self-reliance	0.63
Broadmindedness	0.58

interrelated in such a high degree. These high indices are more likely produced to some extent by what is known among psychologists as the "halo" influence; that is, the tendency for an individual to judge the specific details of another's character according to his general impression of him. This is presumably more likely to be the case where the rating is all done at one sitting rather than in several periods separated from each other by considerable intervals of time.

The reliability of the judgments in this group has been calculated by the well-known method of correlating the combined ratings of one half of the group with the combined ratings of the other half. The table (table 5) showing the coefficients of

reliability for the different traits suggests that our results are entirely satisfactory in this respect.

PERSONALITY AND GENERAL INTELLIGENCE

Common observation would lead one to expect that a good personality would involve at least a fair degree of general intelligence and that a high type of general intelligence would be more likely to be found in a person with a good, than in one with a poor personality. It will be recalled that general intelligence as estimated by one's associates, and personality are very closely related, the coefficient being 0.88; but the correlation in this case is, no doubt, affected favorably by the fact

TABLE 6

Showing the correlations between personality, general intelligence, and school grades for groups I and II

GROUP	CASES	PERSONALITY AND INTELLIGENCE	PERSONALITY AND SCHOOL GRADES	INTELLIGENCE AND SCHOOL GRADES
I	29	0.51	0.61	0.56
II	25	0.34	0.57	0.45

that general intelligence was estimated in connection with the rating of other personal traits. It has already been suggested that the "halo" influence tends to increase the correlations between desirable traits under such conditions. Table 6 showing the correlations between an objective measure of general intelligence, the Army Alpha Test, and personality in these two groups undoubtedly gives us a more accurate notion of the tendency of these two qualities to accompany each other.

SCHOOL GRADES

Reference to table 6 indicates that there is slightly closer relationship between the student's personality and his grades than between his intelligence (measured objectively, not estimated) and his grades. Does this mean that those students who are sometimes accused of deliberately trying to "work"

the instructor for a good grade succeed? To ascertain whether this is the case or not we have correlated the various personal traits with school grades for group I. This will indicate which traits are most effective in the production of high grades assuming that the traits do have a causal relationship to the grades.

If we take those traits showing correlation of 0.50 or better as being of considerable importance, those between 0.30 and 0.49 as having a moderate influence and those below 0.30 as being negligible or of slight importance in the matter of grades, we may then say that, in so far as grades are influenced at all by personal traits, the relative influence of such traits is substantially as follows:

Considerable influence		*Moderate influence*		*Negligible or slight influence*	
Reasoning ability	0.78	Accuracy	0.45	Popularity	0.26
Speed in work	0.73	Originality	0.43	Neatness	0.19
Motor ability	0.59	General ability	0.43	Coöperativeness	0.19
Self-reliance	0.59	Aggressiveness	0.42	Sympathy	0.13
General information	0.57	Reliability	0.41	Appreciation of humor	0.00
Industry	0.56	Enthusiasm	0.34		
Sincerity	0.50	Moral habits	0.33		
		Memory	0.32		
		Address	0.32		
		Social and civic interest	0.31		
		Tact	0.31		

This array of correlations presents splendid evidence that the instructor is not very susceptible in his grading to the diplomatic endeavors of students. The student may laugh ever so heartily at the instructor's stale jokes, he may be ever so sympathetic with his pet theories, and he may present an immaculate appearance in class; but apparently all this will avail him naught so far as grades are concerned if he hath not reasoning ability, mental agility, self-reliance, a good fund of general information, and industrious spirit, and some honesty of purpose. There are, no doubt, instructors who do permit irrelevant personal factors to enter into their grading, but the probabilities seem to be that such instructors are rare as compared with those who disregard all those factors which do not make for substantial scholastic attainment.

How far these results would be valid for other types of schools is problematical. Engineering subjects are of such a nature generally hat achievement in them can be graded rather objectively and accurately. It is likely that the student finds it much easier to "work" his instructor in subjects which involve a good deal of philosophizing and theorizing than in those which require mathematical precision. The writer is inclined to think that such attempts to influence instructors are much less frequent than they are supposed to be and are confined largely to the inferior students.

Having made our analysis of these two groups we shall next attempt to discover how the performances of the men in real life correspond with our various estimates and measures; for after all, the manner in which one stands the tests in the field, in the shop, and in the office will always be regarded as the real criterion of his ability. The measurement, however, of an individual's success under the infinitely complex conditions of business and industrial life presents well nigh unsurmountable difficulties. What progress we have been able to make in this direction will be described in the following section.

(*To be continued*)

GROUP DIFFERENCES IN INTELLIGENCE TESTS

The Relative Difficulty of Types of Questions

GRACE E. MANSON

The University of Michigan

INTRODUCTION

Group intelligence tests can be made to yield more information about individuals and groups than is derived from the final scores. As psychologists, we have overlooked a good deal of valuable information because we have not paused to scrutinize the examinee's reactions to the test situation in detail.

Of course it is valuable to know that a group of expert stenographers makes, on the average, about the same intelligence test score as a group of civil engineers; or that nurses and bookkeepers in the Army were found to be of about the same level and range of intelligence. But the tests can tell us more than this. Indeed, we are more interested in finding out how the nurses differ from the bookkeepers than in finding out the respects in which they are alike. This paper reports some differences—other than differences in intelligence test scores—between two groups of graduate students, a group of industrial engineering students, a group of mechanical engineering students, a group of men engaged in fitting themselves to be salesmen of life insurance, and other groups consisting of minor executives in a manufacturing company, high school students and policemen. The members of each of these groups have in common some such factor as academic training, vocational interest or occupational pursuit. We have raised the question whether their performance in group intelligence tests

cannot be made to yield detailed information concerning specific abilities and disabilities. Are there significant group differences in the relative difficulty of the several types of tasks comprising the intelligence examination?

The answers to such questions would be of value to both the theoretical psychologists and the vocationalists. The former are interested in knowing more concerning the general laws governing interests, motives, attitudes, effects of training, ability to profit by experience, and a host of allied problems; all of which laws are discovered and verified from observing many selected groups in many controlled situations. The vocationalists have to be more specific, and are anxious to find means of ascertaining the interests, motives, attitudes, training, capabilities, etc., of their clients, as well as the methods of putting these forces to the best use when once analyzed. For example, on the basis of final scores, the army psychologists,[1] also Fryer[2] and Burr,[3] have been able to determine the minimum and the maximum and the range of intelligence test scores required for many occupations. But they have also found that the range of test scores at any occupational level is so wide that it includes scores common to many other occupations. It, therefore, becomes necessary to both subject the intelligence examination to more detailed analysis and to devise means for supplementing the intelligence tests by measures of personality and character.

We have limited ourselves to but one phase of the former task, namely, to find the relative difficulty of the several types of questions for groups of different interests and educational status.

[1] Army Mental Tests, Methods, Typical Results and Practical Applications, November 22, 1918, Washington, D. C

[2] Fryer, Douglas, Occupational-Intelligence Standards, School and Society, 1922, Vol. 15, pp. 273-277.

[3] Burr, Emily T., Minimum Intellectual Levels of Accomplishment in Industry, Journal of Personnel Research, 1924, Vol. 3, No. 6, pp. 207-212.

PREVIOUS INVESTIGATIONS

Cobb and Yerkes[4] have been interested in the same problem, with special reference to army officers. They compared the median percentile score made by officers in each branch of the service with the standard score made by all branches taken together (considered the 50 percentile) and plotted psychographs showing the success attained by officers in each group. Their results lead them to conclude that "the psychograph (curve representing measurements for eight types of test which constitute army group-examination) for medical officers differs strikingly from that for artillery officers or from that for engineers. The psychographs of the several special medical groups have a very obvious family resemblance."

Bingham and Davis[5] have carried the investigation into the realm of industry. They studied the intelligence test records of 102 business men attending the annual statistical conference held at the Babson Institute, August 4, 1922. The personal history records of seventy-three of these men were sufficiently complete to classify them in respect to present occupation as business executives, salesmen or doubtful cases. The educational record of each of these men was also known.

The intelligence examination used was not Army Alpha but an adaptation of it known as Bureau of Personnel Research Test VI. Eighty-three per cent of these business men made scores on Test VI which, transmuted into Army Alpha scores, gave them a rating of A or B. The median scores of the three occupational groups were about the same. Examination of the personal history records showed that these men were, on the whole, superior business men. We have here a group of successful executives and salesmen of superior in-

[4] Cobb, M. V., and Yerkes, R. M., Intellectual and Educational Status of the Medical Profession as Represented in the United States Army, Bulletin of the National Research Council, 1921, Vol. 1, Part 8, No. 8, pp. 458–532.

[5] Bingham, W. V., and Davis, W. T., Intelligence Test Scores and Business Success, The Journal of Applied Psychology, 1924, Vol. 8, No. 1, pp. 1–22.

telligence. In these two respects they are very much alike. What does an analysis of the scores made on the six parts of the intelligence examination show? In their conclusions, the authors state:

> The salesmen, as a group, were more proficient upon disarranged sentence questions than were the executives. The executives, on the other hand, were more proficient upon arithmetic and number completion questions. Proficiency on the arithmetic questions shows a negative correlation with schooling beyond grammar school. Proficiency in opposites differentiates the occupational groups not at all but shows some correlation with schooling.

PROBLEM

In the present paper, the writer reports on the relative difficulty of the several types of questions on Bureau of Personnel Research Test VI, as shown by the scores made by eight groups of persons of different ages, educational status and occupational interests. It has been our purpose to see if groups having some factor in common, as education or occupational interest, would find the same types of questions difficult.

SUBJECTS

The test records of 341 persons were studied. These persons were divided into eight groups, as follows:

Group I. Retail bureau students

This group consists of 14 persons, 13 women and 1 man. All of these were college graduates, 10 of them candidates for their Master's degree at the University of Pittsburgh. All of them were interested in preparing to hold executive positions in the personnel department in retail stores.

Group II. Industries students

This group is composed of 28 seniors in the College of Industries at Carnegie Institute of Technology. They were all men in the twenties or thirties, interested in becoming executives or minor executives in the trades of printing, plumbing, building, or machine production.

Group III. Mechanical engineers

Thirty-six seniors in the school of mechanical engineering at Case School of Applied Science make up this group.

Group IV. Carnegie Life Insurance School students

This group includes 36 men who attended the Carnegie Life Insurance School. Many of them were either college graduates or else had attended college. Some of these men had already sold insurance and were sent to the school by their companies, others were taking the course preparatory to selling life insurance. All of them had sufficient interest in life insurance as a vocation to undertake the eleven weeks intensive course.

Group V. Manufacturing company office executives

This group consists of 40 members of the "office" group of a large company manufacturing heating appliances. These men were in the offices of the company in minor-executive positions such as inspectors, supervisors and assistant managers. They were not actively selling but were interested in sales, production and service.

Group VI. Policemen

This group contains 81 policemen of various grades up to lieutenant.

Group VII. Education students

This group is composed of 25 members of a course in mental measurements at the University of Michigan Summer School. In the class were both men and women. Most of them had had several years of experience in teaching or school administration.

Group VIII. High school students

Eighty-one high school sophomores, both boys and girls, make up this group.

THE INTELLIGENCE TEST USED

We have used the same intelligence examination as was used by Bingham and Davis.[6] Bureau of Personnel Research Test VI is a modified Army Alpha in which the questions on arithmetic, opposites, analogies, true-false (disarranged sentences), information and number series are arranged in a

[6] Ibid., pp. 8–9. Also described in Bingham, W. V., Some Group Tests, Psychological Bulletin, February, 1920, p. 57.

seemingly irregular order. The examination is of the spiral omnibus form, each question being slightly more difficult than the question of the same type which preceded it. Correct answers are not weighted but those requiring more time, as arithmetic, number series and true-false, are fewer in number: 20, 20, and 24 as against 40, 40 and 40 for opposites, analogies and information. True-false and opposites, following the army scheme, are scored R-W. Omissions on all tests count as errors but are not doubly penalized in the true-false and opposites tests.

Test VI has been standardized and is published in two forms, Form A and Form B. We have used Form A. The usual time allowed is four minutes for reading the directions and fifteen minutes for the test proper. Fifteen minutes has been our time limit in all groups. However, Hansen and Ream[7] have found that the time allowed did not materially change the standing of persons within a group. Total score on the complete test correlated 0.88 with score at end of 5 minutes, 0.92 with score at end of 10 minutes, 0.96 with score at end of 15 minutes, and 0.97 with score at end of 20 minutes.

METHOD

In determining the relative difficulty of the six types of questions, account has been taken, (1) of the speed factor as shown in the number of problems covered, (2) of the inability factor as shown by the number of examples omitted and in error.

The statistical measure used has, in general, been the average percentage of the group. The average in each case has meant the arithmetic mean. The procedure followed in handling the test records was as here described. Look first at number of the last question attempted. From the key, find out how many arithmetic, analogies, opposites, etc., have been included. Count how many errors have been made on each

[7] Hansen, C. F., and Ream, M. J., The Predictive Value of Short Intelligence Tests, Journal of Applied Psychology, June, 1921, Vol. 5, pp. 184–186.

of the six types of question. Find the number of omissions on each type of question. Figure the per cent of omissions on each type of question by dividing the number of questions of that type included up through the number of the last question the subject attempted, into the number of that type omitted. Figure per cent of error on each type of question by dividing number included (same denominator as for omissions) into number of errors for that type. In the case of true-false and opposites, multiply the number of errors by two before dividing.

TABLE 1

Average scores on Test VI

GROUP	AVERAGE SCORE	
I	124.2	Retail Bureau
II	111.6	Carnegie Industries
III	112.5	Case Mechanical Engineers
IV	105.0	Carnegie Life Insurance Salesmen
V	94.8	Manufacturing Company Minor-executives
VI	38.5	Policemen
VII	121.6	Michigan Summer School 1924
VIII	75.0	High School Sophomores

There were, then, three measures for each subject, (1) number of examples covered, (2) percentage of each type of example omitted, (3) percentage of each type of example wrong. Besides this there was, for each of the eight groups, (1) the average number of examples attempted, (2) the average per cent of omissions on each type of question, and (3) the average per cent of errors on each type of question.

The final score for each subject, using the usual scoring method for Test VI, had previously been determined, as well as the average score made by each group. These average final scores are given in table 1.

Group I and Group VII have approximately the same average score. These are the Retail Bureau and the Michigan Summer Session groups. Both groups are composed of many

college graduates. Groups II and III have average scores of about the same size. These are the two applied science groups, industries and mechanical engineering. Group IV is the Life Insurance group, Group V the lesser executives group. Group VIII, the high school group, is next lowest with Group VI, the policemen, as the most inferior group.

These means are in keeping with the norms found for comparable groups with Test VI. The life insurance group here is above the expected norm having a mean of 105, whereas a mean of 105 falls in the eighth decile for a group of 537 members of the Carnegie Life Insurance School. The norms for 100 college seniors place scores of 111 and 112 in the eighth decile. But since engineering students are possibly a more selected group, these means are not particularly high. Scores of 121 and 124 are in the ninth decile for college seniors which would be about where we would expect to find graduate student groups like Group I and Group VII. A mean of 75 puts the high school group in the second decile of college seniors which is about what we would anticipate. The minor-executive group with a mean of 94.8 is to be compared with a mean of 107 found for salesmanagers at a Babson conference and of 75 for salesmen belonging to the same company as these minor-executives. The policemen are much lower than the high school students and come in the second decile of a group of order taking salesmen, which is probably the easiest type of selling.

As a whole we should say that these groups are representative of their classes with the exception of the life insurance salesmen who have scores somewhat too high.

THE SPEED FACTOR

Since all groups were given the tests under the same time conditions, fifteen minutes, a comparison of the number of the last example attempted would give some evidence of differences in the speed factor.

The difference between the number of the last example attempted and the final score is equal to the sum of the errors

and the omissions. The size of these differences follows the inverse order of the number of the last example attempted. Groups which have attempted the most examples have made the smaller number of errors and omissions. This is in keeping with previous experimental findings that speed and accuracy are highly correlated. The life insurance salesmen and the minor-executives change places in this relationship, the insurance group has made more errors and omissions than expected, the minor-executives group less.

Test VI is so arranged that it is impossible to determine the speed factor for each type of question. However, the

TABLE 2

Average number of last example attempted

	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	GROUP VII	GROUP VIII
Number of last example attempted	140.5	127.5	130.2	126.6	112.4	67.5	137.8	99.2
Final score	124.2	111.3	112.5	105.0	94.8	38.5	121.6	75.0
Difference	16.3	16.2	17.7	21.5	17.6	29.0	16.2	24.2

general conclusion can easily be made that the difficulty of the questions for the group is a determining factor since groups which have covered little ground have also made many errors and omissions.

RELATION OF NUMBER OF OMISSIONS TO NUMBER OF ERRORS

These facts are quite clearly brought out in the next table which shows the relation of the number of omissions to the number of errors for each group, and the relation of each of these to the number attempted (table 3).

Table 3 gives the gross averages of errors and omissions. The total number attempted is not taken into account. This relationship will be shown in percentage form in a later table (table 4). The interesting points to note are:

1. That the average number of omissions is high where the average of the last attempted is low, Groups V, VI and VIII.

2. That where the average number of last attempted is low, the ratio of errors to omissions is high. Group I is an exception to this.

TABLE 3

Average number of errors and omissions

	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	GROUP VII	GROUP VIII
Omissions	3.8	1.2	1.3	1.9	5.6	11.0	2.2	5.4
Errors	12.5	15.0	16.4	19.7	12.0	18.0	14.0	19.0
Number last attempted	140.5	127.5	130.2	126.6	112.4	67.5	137.8	99.2
Final score	124.2	111.3	112.5	105.0	94.8	38.5	121.6	74.8

TABLE 4

Average percentages of averages of errors and omissions

	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	GROUP VII	GROUP VIII
Omissions	0.027	0.009	0.009	0.015	0.05	0.165	0.016	0.054
Errors	0.081	0.117	0.126	0.156	0.107	0.266	0.101	0.191
Number last attempted	140.5	127.5	130.2	126.6	112.4	67.5	137.8	99.2

3. That groups of approximately the same speed show individual differences in their choice between omitting and making errors, Groups I and VII.

Table 4 shows these same relationships in percentage form. This per cent of error means per cent of total number covered (number of last question attempted) that are wrong. Per cent omitted means per cent of total number covered (number of last question attempted) that are omitted. Thus, account is taken of the number covered when considering the number omitted and wrong.

This table brings out very clearly several points:

1. That, on the whole, knowledge of the material governs the speed factor, which is shown by the number of examples covered, since percentages of both errors and omissions are low where speed is high. See Groups I, II, III and VI.

2. That when the examples become very difficult for the group, the number of omissions and the number of errors both increase, the ratio between the per cent of omissions and per cent of errors approaching a 1:1 condition. A 1:2 ratio is found for Groups V and VI.

3. That groups show individual differences in choosing between omissions and errors, showing that it is likely that factors of "caution"[8] and "rashness" are operative. Groups IV and V are noteworthy examples of this. It is questionable whether this urge, to be cautious or rash, becomes operative to any degree until the material reaches a certain degree of difficulty. A more detailed discussion of this relationship will be found in a later paper.[9]

RELATIVE DIFFICULTY OF THE TESTS AS SHOWN BY OMISSIONS

Another point of interest is to determine whether the high percentages of errors and omissions concentrate on the same type of questions in all groups. In the next table, table 5, is given the percentages of omissions of the total number tried for each group, arranged in the order of increasing amounts. It should be noted that this table shows relative difficulty within each group, not between groups.

If we study the concentration of omissions by comparing the five groups making the highest scores (all above 100), Groups I, II, III, IV and VII, with the three groups making the lowest score (all below 100), Groups V, VI and VIII, certain interesting points are brought out.

[8] Brown, Wm., A Study of the "Caution" Factor and Its Importance in Intelligence Test Performance, American Journal of Psychology, 1924, Vol. 35, No. 3, pp. 368-381.

[9] Manson, Grace E., The Significance of "Omissions" on a Group Intelligence Test.

1. The high score groups have their larger percentages of omissions in the number series, information and arithmetic tests, their small percentages of omissions on the true-false, analogies and opposites tests. The small percentage of omissions on information items in Group III is the only exception to this.

2. The groups making lower total scores are not as consistent but have certain points in common. Opposites and true-false remain in the few omissions class but analogies

TABLE 5

Average percentages of averages of omissions on each type of question

GROUP I		GROUP II		GROUP III		GROUP IV	
T. F.	0.000	Opp.	0.000	Inf.	0.001	Anal.	0.000
Anal.	0.0025	T. F.	0.0018	Opp.	0.002	T. F.	0.000
Opp.	0.005	Anal.	0.005	T. F.	0.003	Opp.	0.002
No. S.	0.038	Inf.	0.009	No. S.	0.015	Inf.	0.01
Inf.	0.052	No. S.	0.013	Anal.	0.018	No. S.	0.033
Arth.	0.09	Arth.	0.047	Arth.	0.028	Arth.	0.035

GROUP V		GROUP VI		GROUP VII		GROUP VIII	
No. S.	0.012	Arth.	0.03	Anal.	0.000	T. F.	0.01
Opp.	0.024	Opp.	0.09	T. F.	0.000	Opp.	0.02
T. F.	0.027	T. F.	0.098	Opp.	0.01	No. S.	0.03
Arth.	0.044	No. S.	0.12	Arth.	0.019	Anal.	0.04
Inf.	0.05	Inf.	0.22	No. S.	0.024	Arth.	0.06
Anal.	0.111	Anal.	0.28	Inf.	0.04	Inf.	0.13

moves to the greater omissions class, having the largest number of omissions in Groups V and VI. Arithmetic and information remain in the more omissions class, the same as with superior subjects.

True-false and opposites are consistently among the fewer number of omissions in each of the eight groups. The order is sometimes reversed but in all groups true-false and opposites are found among the three lowest percentages of omissions. This is no doubt due to the fact that the guessing element is made very evident in the form of the questions, the 50-50

chance of being right is reacted to by the majority of subjects. But when the words are very difficult for the subjects, even this lure is ignored, for policemen omitted 9.8 and 9 per cent of the true-false and opposites questions.

On the other hand, arithmetic examples are omitted more often than other types of questions, that is, they fall among the highest three percentages for all groups except the policemen. Number series has somewhat the same tendency although not as decided. These two types of tests are the least subject to guessing and also take more writing. Thus, for persons interested in speeding up, these would be logical places to omit.

Analogies and information items offer a 1:4 guessing proposition. If realization of the chance element were operative alone, we should expect analogies and information items to stand third and fourth in the rank of omissions. But from their positions in the table, it seems likely that the difficulty of the operations had more influence. For analogies are omitted more often than any other items by the policemen and the minor-executives groups and least often by the life insurance and Michigan graduates groups. Information test percentages are irregular in position. In five out of the eight groups, they are ranked last or next to last. Difficulty therefore seems to be the stronger force.

Thus we might conclude that the percentage of omissions is controlled, (1), by the form of the question, that is, likelihood of success in guessing; (2), by the difficulty of the type of question; (3), by the time that the type of question required in solving; (4), by the operating strength of the directions "not to omit;" and (5) by other reasons particular to the group or to individuals in the group.[10]

RELATIVE DIFFICULTY OF THE TESTS AS SHOWN BY ERRORS

In this section the relative difficulty of the several types of questions will be shown by the percentages of error. The

[10] Supporting evidence gained through introspective reports is given in the article previously referred to.

chance factor of success on true-false and opposites was eliminated by the scoring used; each error deducted two from the total score. The percentages of error are shown in table 6.

It is to be noted that opposites fall among the three items having the largest percentages of error in three of the superior groups, I, II and III, and rank third in Groups IV and VII. True-false ranks second in Groups I, II and VII, fourth and

TABLE 6

Average percentages of averages of error on each type of question

GROUP I	GROUP II	GROUP III	GROUP IV
Anal. 0.04	Anal. 0.07	Arth. 0.077	Anal. 0.11
T. F. 0.068	T. F. 0.096	No. S. 0.079	Arth. 0.12
No. S. 0.07	No. S. 0.11	Anal. 0.08	Opp. 0.134
Arth. 0.09	Arth. 0.114	T. F. 0.088	No. S. 0.15
Opp. 0.10	Opp. 0.118	Opp. 0.136	Inf. 0.18
Inf. 0.16	Inf. 0.19	Inf. 0.23	T. F. 0.228

GROUP V	GROUP VI	GROUP VII	GROUP VIII
Arth. 0.066	Arth. 0.076	Anal. 0.03	Anal. 0.11
Opp. 0.09	Inf. 0.18	T. F. 0.054	No. S. 0.12
Inf. 0.10	No. S. 0.20	Opp. 0.10	T. F. 0.146
No. S. 0.12	Opp. 0.30	No. S. 0.13	Arth. 0.15
Anal. 0.12	Anal. 0.35	Arth. 0.16	Inf. 0.22
T. F. 0.134	T. F. 0.44	Inf. 0.16	Opp. 0.28

sixth in Groups III and IV. In the three low score groups, opposites rank respectively 2, 4 and 6, true-false as 6, 6 and 3. Either true-false or opposites rank among the higher percentages of error in all of the low score groups and in all but one of the high score groups. Thus, when the guessing element is penalized in the scoring, true-false and opposites are not as easy as one might expect from looking at omissions alone.

Analogies and information items are scored in the same manner for errors as for omissions. Information items are difficult for the superior groups, being the most difficult type of question for four out of five of these groups and next to the most difficult for the other one. In the case of the low score

groups, the information items rank 3, 2 and 5. In relation to the other types of questions, the information questions are harder for the high score groups than for the low score groups. The reverse condition is true for analogies. Analogies rank 1, 1, 3, 1 and 1 for the high score groups, and 5, 5 and 1 for the low score groups.

The tests dealing with numbers, arithmetic problems and number series, are probably the most time consuming of all the tests. Nevertheless, arithmetic problems are the easiest kind of questions for the minor-executives and policemen, and hold an intermediate position for the high school group. In the high score groups, arithmetic problems fall in the medium percentage groups, ranking 4, 4, 1, 2 and 5. For two of the low score groups, arithmetic problems are easiest. For three of the high score groups, they are among the more difficult. Number series ranks about third with the high score groups, standing 3, 3, 2, 4 and 2. In the low score groups, number series stand 4, 4 and 1. We may then conclude that the arithmetic questions are easier than the other types of questions for the low score groups and among the harder types of questions for the high score groups. Number series are slightly easier for the high score groups than for the low score groups. On the whole, however, we may say that on the basis of errors alone, arithmetic and information are among the hardest tests for the high score groups and among the easier tests for the low score groups. The number series test takes an intermediate position. This conclusion is not without exceptions, the Case group and the high school group not following the order of their classification.

RELATIVE DIFFICULTY OF THE TESTS AS SHOWN BY BOTH ERRORS AND OMISSIONS

If we combine errors and omissions, certain very interesting contrasts stand out, as shown in table 7.

The most important point that this table brings out is that the high score groups find the word abstractions, like analogies, true-false and opposites, the easier type of questions,

whereas the low score groups find them the more difficult types. Policemen and minor executives find the arithmetic, number series, and information items the easier. These are more concrete types of question, dealing very much less with the finer, more exact knowledge of words and written language.

Interesting exceptions to the general tendency are found in Group III, the mechanical engineering group, and Group VIII, the high school group. In the Case engineering group,

TABLE 7

Average percentages of averages of errors and omissions combined

GROUP I	GROUP II	GROUP III	GROUP IV
Anal. 0.0425	Anal. 0.075	T. F. 0.091	Anal. 0.11
T. F. 0.068	T. F. 0.0978	No. S. 0.094	Opp. 0.136
Opp. 0.105	Opp. 0.118	Anal. 0.098	T. F. 0.155
No. S. 0.108	No. S. 0.123	Arth. 0.105	Arth. 0.155
Arth. 0.18	Arth. 0.161	Opp. 0.138	No. S. 0.183
Inf. 0.212	Inf. 0.199	Inf. 0.231	Inf. 0.19

GROUP V	GROUP VI	GROUP VII	GROUP VIII
Arth. 0.110	Arth. 0.106	Anal. 0.03	No. S. 0.15
No. S. 0.132	No. S. 0.32	T. F. 0.054	Anal. 0.15
Inf. 0.15	Opp. 0.39	Opp. 0.11	T. F. 0.156
Opp. 0.158	Inf. 0.40	No. S. 0.154	Arth. 0.21
T. F. 0.161	T. F. 0.538	Arth. 0.179	Opp. 0.30
Anal. 0.245	Anal. 0.63	Inf. 0.20	Inf. 0.35

success in opposites, an ability dependent upon word knowledge, drops below number series and arithmetic. Since mathematics is stressed in the engineering schools, rather than the cultural subjects, this condition seems logical. The other possibility is that persons interested in figures rather than niceties of words go into the engineering schools. The high school group, a mixed and undeveloped group, is irregular, neither the word nor the number problems predominating.

Looked at from another point of view, the school and the business groups, we find that the five school groups, I, II, III, VII and VIII are better at handling word relations than at

handling arithmetic and general information items, whereas the outside groups, IV, V and VI are better at solving questions of general information and arithmetic. Group IV, the life insurance group, follows the rule of the school groups. This is likely due to the fact that it is a Life Insurance School group. The high school group, although more like the school group, has some of the characteristics of the "outside groups." The mechanical engineers have some of the characteristics of the school groups, some of the "outside groups."

As a general conclusion, we might say that the academic groups have been more successful in handling abstract word relations than in handling items of number relation and general information, and that the "outside groups" have been more successful in solving arithmetical relations and answering questions of general information than in solving problems which involve the handling of abstract word relations. This tendency is particularly strong in the case of analogies and arithmetic problems. The impression should not be gained, however, that the academic groups cannot do arithmetic problems or that the outside groups are apt at them. The mean score of Group I (high score) for arithmetic problems is, for example, 12.5 whereas for Group VI (low score) the mean score is 6.67, yet the first group has made a lower mean score on arithmetic than on any of the other types of questions and the other group has made a higher mean score on arithmetic than on any of the other questions.

Reasons for these differences in the relative difficulty of the several types of problems are more in the form of conjectures and hypotheses than in the realm of established facts. Immediate interests, which are always a governing factor, may have caused one group to work one type of question with more ease than another type. The familiarity of the content and methodology is another likely cause of differentiation.

The influence of interests makes one wonder if a change of environment would change the type of question answered most successfully. Would a group of college graduates, who had become successful business executives, excel in arithmetic and

general information rather than in analogies, opposites or true-false examples? Bingham and Davis[11] have taken up this problem, considering a group of successful salesmen and executives who had taken Bureau Test VI. They report: "Proficiency in arithmetic questions shows a negative correlation with schooling beyond grammar school." "Proficiency in opposites differentiates the occupational groups not at all, but shows some correlation with schooling." There then seems to be some evidence that this greater interest and facility in the abstract holds after the college period.

SUMMARY

Are there group differences in the relative difficulty of the several types of tests comprising an intelligence examination?

To answer this question, we have studied the mental test records of 341 persons, comprising eight groups of different occupational and educational interests. The relative difficulty of the six types of questions has been determined from the number of examples attempted, the number of errors and the number of omissions. Although the number of cases included and the differences between steps in difficulty are not great, certain tendencies are sufficiently outstanding to be of some significance. The value of some of these findings is further enhanced by the fact that they verify the results of Cobb and Yerkes, and Bingham and Davis. In comparison with these investigators, our results show:

1. That seniors in mechanical engineering find opposites and information items the most difficult types of question; number series and arithmetic problems among the easier problems. The order of difficulty from easiest to hardest is: (*a*) True-false, (*b*) Number Series, (*c*) Analogies, (*d*) Arithmetic, (*e*) Opposites, (*f*) Information. Cobb and Yerkes,[12] with a very much larger group including men in all branches of en-

[11] Bingham, W. V., and Davis, W. T., Intelligence Test Scores and Business Success, Journal of Applied Psychology, 1924, Vol. 8, No. 1, pp. 1–22.

[12] Op. cit., p. 478.

gineering, found the order of increasing difficulty to be: (*a*) Number Series, (*b*) Analogies, (*c*) Arithmetic, (*d*) True-false, (*e*) Information, (*f*) Opposites.

2. That insurance salesmen, Group IV, find analogies, opposites and true-false questions easier than arithmetic, number series and information items. The order of increasing difficulty is as given in the preceding sentence. Bingham and Davis[13] report: "Thus, we suggest that disarranged sentences, arithmetic and number completion problems be included in tests designed to differentiate prospective salesmen from other business men, the hypothesis being that superiority in disarranged sentences, coupled with inferiority in arithmetic and number completion, is more often typical of salesmen."

3. That minor-executives, Group V, find arithmetic, number series and information problems among the easier types, with opposites, true-false and analogies as the more difficult types. Bingham and Davis[13] found "The executives, on the other hand, were more proficient upon arithmetic and number completion questions."

From the point of view of educational status, we have found:

1. That groups of graduate students, whose college courses have been in the Arts College or in Schools of Education, find the order of increasing difficulty to be as follows: (*a*) Analogies, (*b*) True-false, (*c*) Opposites, (*d*) Number Series, (*e*) Arithmetic, (*f*) Information. This is the order for Groups I and VII. Group II, which is composed of industries students, follows the same order. In this group, the academic influence seems to be stronger than the influence of the vocational interest.

2. That mechanical engineering students follow more closely the order for engineers than for graduate students specializing in the more philosophical subjects.

3. That high school students, Group VIII, find opposites difficult and number series easy, which makes them more like the mechanical engineering students than like the Arts students.

4. That, as a whole, the "in-school" groups, I, II, III, IV,

[13] Op. cit., p. 21.

VII and VIII, find the abstract word relation items, analogies, true-false and opposites, easier than the more concrete examples, number series, arithmetic and information. This fact is brought out particularly by the comparatively large percentages of errors in arithmetic and information items. On the other hand, the two "out of school" groups, V and VI, find true-false and analogies to be very difficult items and arithmetic and number series comparatively easy items.

The fact that the information items are the most difficult ones for all the school groups seems to give some verification to the opinions of Mr. Edison and others, that our college graduates do not have a wealth of specific information.

It is our belief that information of value has been gained from the more detailed analysis of mental test records. However, this information must be verified by several independent investigators before it can be used in the service of vocational guidance or industry.

TWO STUDIES OF ATTENTION TO ADVERTISEMENTS

H. K. NIXON

Psychological Laboratory, Columbia University

INTRODUCTION

In this article are reported two studies of attention to advertisements, one having to do with the influence of color and the other dealing with the influence of pictures of people in an advertisement. The data here presented are derived from the results of a research previously published by the author[1] and are intended as an extension of the conclusions of that investigation.

The experiment from which the present results were secured may be briefly described. In general it consisted in presenting two advertisements simultaneously, one advertisement possessing the feature whose attention value was being tested, and the other not possessing this factor, and observing the length of time the subject looked at each advertisement. The full technique of this experimental method and results showing its validity have already been presented[2] and will not be discussed here. Sufficient to say that it serves as an objective approach to problems of attention which have previously been attacked by less direct methods.

The subjects of the experiment were 30 in number, consisting of business men, housewives, and students in Columbia University, 14 men and 16 women. The age range was from eighteen to fifty. In the first experiment here reported thirty-four advertisements having pictures of people and thirty-four having pictures of objects were paired. In the experiment on color

[1] H. K. Nixon, Attention and Interest in Advertising, Archives of Psychology, Columbia University, R. S. Woodworth, Editor, 1924.

[2] Loc. cit.

forty-four colored and forty-four black and white advertisements were paired. In the first case the factors of position (right or left hand page), color, size of illustration, possession of border, and degree of complexity were kept in balance, so that the pictures of people appeared half of the time on the left hand page, half of the time with color, and so on. In the case of the color experiment similar care was taken. All of the advertisements used were full page, of Saturday Evening Post size. They were bound together in a book of which the subject turned the pages. For the part of the experiment dealt with here the subject spent thirty seconds on each pair, distributing his time between the members of the pair as he was inclined. Unknown to the subject the observer recorded the distribution of eye fixations and the following pages present the tabulation of the data thus derived in the case of pictures of people and of color.

I. THE USE OF PICTURES OF PEOPLE AS AN ATTENTION FACTOR IN ADVERTISING

The practical advertising man has always recognized the value of illustrations of people as a method of attracting attention. A survey of current magazines shows that over 65 per cent of full page advertisements contain pictures of humans as a prominent feature, and it is probable that this use is on the increase. Academic psychology has likewise noted that people are prepotent factors in attention. As far as the writer knows, no study has ever been made to determine with exactitude the real efficiency of illustrations of people. Hollingworth[3] has, indeed, cited an experiment showing that persons and faces are more easily remembered than objects and he has stressed the value of suggested activity, which is likely to be the activity of humans. Aside from this, the topic seems to have been neglected. As the data here presented will indicate, this neglect is serious, for in any experiment on advertisements one might almost as well neglect size or position and might better neglect

[3] Hollingworth, H. L., Advertising and Selling, p. 114.

color or border or size of display type than to neglect this factor. It is, therefore, both from the point of view of the advertiser and that of the experimenter on advertising material, of importance that definite data be secured on this subject. The figures presented in this article are suggestive of what may develop.

The data of the original experiment of which this work is an elaboration indicated that subjects, in a surprising large number of cases, tend to first turn the eyes to advertisements having pictures of people in them, as opposed to advertisements having only illustrations of objects. For convenience the original figures are reproduced in table 1.

The original data also showed that pictures of people maintained a decided interest throughout the period of exposure.

TABLE 1

First fixations out of possible 34 (Thirty subjects)

	AVERAGE NUMBER	$S.D._{Av.}$
34 advertisements with pictures of people...	18.93	0.48
34 advertisements with pictures of objects...	15.07	0.48

Difference = 3.86. $S.D._{Dif}... = 0.96$.

Averages were worked out showing the accumulated attention value at the end of ten seconds and at the end of thirty seconds. As a general summary these averages were valuable but it seemed desirable to have a more detailed picture of the course of attention. For this purpose the data are so tabulated as to show how much time pictures of people received during each two-and-a-half-second interval throughout the thirty seconds. This is shown in table 2 and is represented graphically in figure 1. Since in the experiment the half-second was used as the unit of time recording all of the figures are in half-seconds.

A consideration of table 2 shows clearly the efficiency of pictures of people as attention factors. In the columns numbered 1 to 12 are given the average times spent by the 30 subjects on the 34 advertisements having pictures of people, cal-

culated for each five half-seconds during the total thirty second exposure. By the method of presentation the group having pictures of objects and the group having the pictures of people had an equal chance to attract attention and had there been no difference in their attention value the averages in each five half-second interval would be 2.50 for each. It will be noted, however, that in most instances the values in table 2 are considerably greater than this, and the averages of the averages range from 2.77 to 3.14. This means that at no time during the thirty seconds were pictures of objects as efficient as their rivals. The small S.D.'s of these averages indicate that they are reliable and by these figures there is not one chance in ten thousand that in any interval the true average would fall as low as 2.50, the point at which the two kinds of illustrations would be equal. It may be said that these results are substantiated by two other experiments which will not be discussed here, in which other groups of subjects and other material gave averages even more favorable to pictures of people.

In figure 1 these facts are presented in graphic form. It will be noted that pictures of people begin with a superior record for first fixations (FF), the curve rising for about twenty-five half-seconds. This indicates that not only does this device attract immediate attention but it serves to hold interest and in this respect is probably superior to any other yet tested. This first period of rise of the curve is followed by a gradual fall, the curve never approaching, however, in the total thirty seconds very much nearer the mid-line than it was at first. If we adopt the somewhat misleading graphic expedient of allowing the area below the curve to represent relative attention to advertisements having pictures of people, the superiority of this group is evident to the eye.

From the original experiment there are also available data showing the memory value of advertisements having pictures of people, as derived from a recall memory test. That value is indicated on the figure by the column at the right. From this it is apparent that the superior attention getting power of these advertisements is reflected in greater frequency of recall.

TABLE 2

Average time spent on 34 advertisements with pictures of people during each 2½ second interval for 30 seconds

(Figures are in half-seconds)

SUBJECT NUMBER	AVERAGE TIME OUT OF POSSIBLE 5 HALF-SECONDS											
	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10	Interval 11	Interval 12
1	2.38	3.00	3.11	2.94	3.32	3.14	2.79	3.32	3.52	3.11	3.02	2.64
2	2.91	3.17	3.02	3.14	3.44	3.14	2.94	3.20	3.20	3.02	3.29	3.32
3	2.79	3.47	3.50	3.41	3.61	3.82	3.61	4.17	4.05	4.20	4.20	4.00
4	2.67	3.05	4.11	3.73	3.47	3.05	2.91	2.55	2.58	2.23	1.73	2.23
5	2.85	2.76	2.58	2.85	2.70	2.52	2.47	2.52	2.41	2.23	2.70	2.58
6	3.23	3.67	3.47	2.94	2.79	2.76	2.41	2.32	2.29	2.26	2.38	2.55
7	2.85	3.67	3.76	3.67	3.52	3.61	3.61	3.38	3.44	3.05	3.29	3.35
8	2.61	2.67	2.79	3.29	3.26	3.05	3.14	2.58	2.85	2.97	3.08	3.02
9	3.20	3.02	3.41	2.94	3.02	3.00	2.55	2.38	1.88	1.82	1.91	2.05
10	2.23	3.11	3.38	3.38	3.23	3.26	3.35	3.17	2.79	2.50	2.73	2.88
11	2.67	2.61	2.52	2.85	2.38	2.73	2.88	2.97	2.64	2.64	2.55	2.64
12	3.11	4.05	4.05	4.02	4.00	3.29	3.20	2.70	2.67	2.29	2.11	2.11
13	3.11	3.35	2.94	2.79	1.97	2.02	2.17	1.94	2.05	2.26	2.97	2.94
14	2.67	2.44	2.23	2.20	2.52	2.58	2.52	2.50	2.26	2.14	2.11	2.14
15	2.85	3.50	3.11	3.38	3.52	3.58	3.14	2.73	2.47	3.23	2.88	2.76
16	2.58	2.61	2.58	2.67	2.26	2.05	2.14	2.38	2.44	2.58	2.70	3.02
17	3.14	3.35	3.88	3.76	4.00	4.00	3.44	3.41	3.23	3.55	3.44	3.20
18	3.44	2.88	2.94	3.26	3.47	3.70	3.44	3.26	3.23	3.23	3.32	3.08
19	2.70	3.44	2.94	2.97	3.32	3.58	3.29	3.26	3.61	3.58	3.52	3.26

20	2.50	2.44	2.35	2.70	2.73	2.94	2.82	2.85	3.11	3.32	3.26	3.05
21	2.97	2.85	2.61	3.41	3.47	3.50	3.50	3.52	3.58	3.52	3.55	3.64
22	3.00	2.32	1.79	2.74	2.70	2.97	3.55	3.35	2.88	2.52	2.58	2.58
23	3.08	2.44	2.91	2.76	3.26	3.47	2.94	2.91	2.55	2.32	2.47	2.14
24	2.47	2.52	2.94	3.20	3.05	2.85	2.70	2.61	2.85	2.88	2.82	2.94
25	2.38	2.64	2.38	2.67	2.91	2.67	2.91	2.91	2.76	2.58	2.64	2.23
26	4.17	4.20	4.32	4.08	4.05	3.79	3.41	2.79	2.79	2.79	2.73	2.82
27	2.50	2.73	2.88	3.23	3.61	2.94	2.44	2.38	2.50	2.67	2.61	2.64
28	2.76	2.70	2.88	2.55	2.82	2.88	3.08	3.38	3.26	2.73	2.73	2.70
29	3.00	3.08	2.61	2.29	2.32	2.76	2.64	2.26	2.38	2.32	2.64	2.32
30	3.67	3.29	3.23	3.32	3.55	3.41	3.38	3.47	3.29	2.79	2.47	2.38
Average ..	2.88	3.03	3.07	3.10	3.14	3.10	2.98	2.90	2.85	2.78	2.81	2.77
$S.D._{Ave.}$	0.073	0.081	0.105	0.084	0.097	0.087	0.077	0.088	0.092	0.095	0.094	0.085

It would be interesting to know the real reason or reasons for the advantages shown by these advertisements. We may

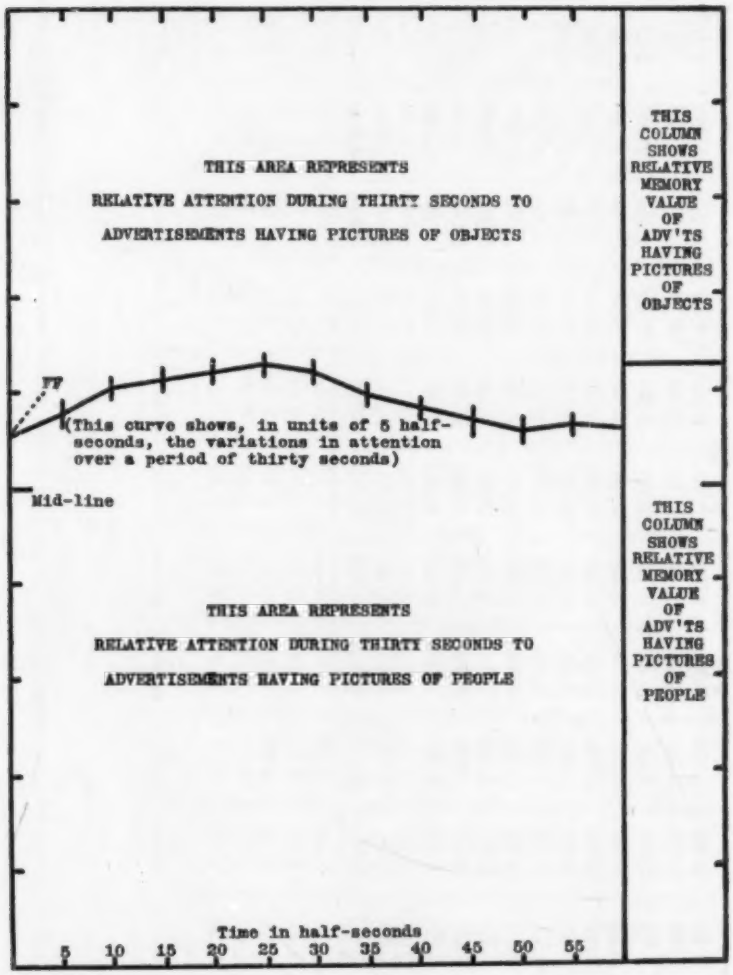

THIS AREA REPRESENTS
RELATIVE ATTENTION DURING THIRTY SECONDS TO
ADVERTISEMENTS HAVING PICTURES OF OBJECTS
THIS COLUMN SHOWS RELATIVE MEMORY VALUE OF ADV'TS HAVING PICTURES OF OBJECTS
FF
(This curve shows, in units of 5 half-seconds, the variations in attention over a period of thirty seconds)
Mid-line
THIS AREA REPRESENTS
RELATIVE ATTENTION DURING THIRTY SECONDS TO
ADVERTISEMENTS HAVING PICTURES OF PEOPLE
THIS COLUMN SHOWS RELATIVE MEMORY VALUE OF ADV'TS HAVING PICTURES OF PEOPLE
Time in half-seconds
5 10 15 20 25 30 35 40 45 50 55

FIG. 1

speculate that it arises from an interest, either innate or acquired as a result of many experiences, in human beings and in all representations of their persons and acts. This is, of course, the

old theory of "human interest." It seems to the author, however, that there is likely to be more involved than that. For one thing, the situations in which people are presented are not so likely to be apparent at a glance. The time taken "to see what it is all about" is likely to be longer and the advertisement thus gains extra credit. It is likely that many other factors come in to influence the final showing. It would be foolish, then, to think of a picture of a person as being comparable in simplicity as an advertising device to color or border or type. Much more experimental work is necessary before we can hope to tell exactly why it is that such a picture is more attractive than one of some common object.

From the practical point of view it should therefore be understood that there is no magic by which an indiscriminate sticking of such pictures of people into an advertisement is going to give that advertisement superior attention power. It is not as simple as putting in color. By a proper use of pictures of humans the advertiser can very probably add considerably to the attractiveness of any advertisement, but what that proper use is remains for further investigation.

II. THE ATTENTION VALUE OF COLOR IN ADVERTISEMENTS

Franken and Hotchkiss[4] found that colored advertisements had a 13 per cent greater recognition memory value than an ordinary black and white page. Kitson[5] reports that the use of color is greatly on the increase since 1912. These two investigations are practically the only ones available on this important practical question of the value of the use of color as an attention device. The impression is widespread that color is very effective and millions of dollars are spent every year under the idea that if the advertisement is only given color it is sure to dominate its black and white competitors.

[4] Hotchkiss, G. B., and Franken, R. B., The Attention Value of Advertisements in a Leading Periodical, published by Graduate School of Business Administration, N. Y. University, 1920.

[5] Kitson, H. D., Minor Studies in the Psychology of Advertising, Journal of Applied Psychology, 6, 1922, pp. 59-68.

TABLE 3

Average time spent on 44 advertisements with color during each 2½ second interval for 30 seconds

(Figures are in half-seconds)

SUBJECT NUMBER	AVERAGE TIME OUT OF POSSIBLE 5 HALF-SECONDS											
	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10	Interval 11	Interval 12
1	3.00	1.90	1.95	2.18	2.02	1.68	1.61	2.02	2.09	2.31	2.38	2.18
2	2.86	2.68	2.65	2.47	2.24	1.90	2.02	2.09	2.06	2.11	2.06	2.09
3	2.58	2.68	2.65	2.54	2.47	2.86	2.65	2.63	2.68	3.02	3.22	3.34
4	2.27	2.31	2.52	2.81	2.72	2.13	2.00	2.00	2.20	1.86	1.93	2.15
5	2.38	1.93	1.86	2.29	1.88	2.15	2.36	2.34	2.45	2.38	2.13	1.97
6	2.65	2.59	2.34	2.24	1.90	1.97	1.93	1.63	1.59	1.81	1.86	2.09
7	2.09	1.70	2.02	1.97	1.93	1.97	1.97	1.93	1.88	2.22	2.45	2.43
8	2.59	2.63	2.59	2.49	2.68	2.59	2.11	2.45	2.22	2.04	2.29	2.49
9	2.49	2.72	3.00	2.29	2.09	1.93	2.11	2.54	2.29	2.09	2.09	2.11
10	2.45	2.47	2.29	2.06	2.29	3.04	2.61	2.31	2.38	2.18	2.20	2.13
11	2.97	2.72	2.84	2.84	2.72	2.74	2.81	2.77	2.84	2.79	2.65	2.81
12	2.36	2.45	2.18	1.95	1.97	1.65	1.63	1.68	1.95	2.59	2.43	2.56
13	1.81	2.04	2.06	2.56	2.24	2.38	2.56	2.49	2.40	2.45	2.77	2.86
14	3.00	2.63	2.40	3.04	3.02	2.88	2.65	2.63	2.88	2.90	2.93	2.97
15	2.93	2.36	2.13	2.18	2.27	2.31	2.20	2.24	2.11	2.36	2.34	2.22
16	2.88	3.09	3.13	2.77	2.15	1.86	1.77	1.77	1.97	1.97	2.00	2.09
17	2.56	2.97	3.09	2.88	3.02	3.06	2.97	2.97	2.95	3.43	3.06	2.79
18	2.79	2.54	2.45	2.52	2.77	2.81	2.81	2.63	2.77	2.45	2.59	2.65
19	3.22	2.63	2.36	2.40	2.38	2.38	2.11	2.00	1.97	2.20	2.13	2.61

20	2.49	2.49	2.49	2.47	2.54	2.29	1.81	2.05	2.15	2.27	2.15	2.11
21	2.38	2.61	2.38	2.36	2.38	2.36	2.20	2.38	2.56	2.72	2.84	3.11
22	2.63	2.65	2.72	2.24	2.54	2.29	2.72	2.70	2.95	2.52	2.61	2.54
23	2.49	2.59	2.54	2.48	2.63	2.65	2.40	2.22	2.20	2.20	2.22	2.31
24	2.72	2.70	2.45	2.24	2.11	2.34	1.68	2.00	1.97	1.95	1.79	1.65
25	3.09	3.02	2.72	2.15	2.22	2.09	2.56	2.72	2.68	3.02	2.84	2.77
26	2.77	2.84	2.74	2.45	2.70	2.54	2.38	2.24	2.34	2.38	2.52	2.59
27	2.56	2.34	2.77	2.68	2.81	2.77	2.93	2.74	2.49	2.36	2.13	2.29
28	2.54	2.54	2.63	2.27	2.77	2.65	2.68	2.34	2.11	2.18	2.45	2.49
29	2.59	2.79	2.93	2.84	2.54	2.36	2.77	2.68	2.68	2.40	2.45	2.27
30	2.56	2.95	2.93	2.72	3.24	3.20	2.81	2.81	2.52	2.86	2.59	2.86
Average ..	2.62	2.55	2.54	2.45	2.44	2.39	2.33	2.33	2.34	2.40	2.40	2.45
$S.D._{Ave}$	0.053	0.058	0.054	0.050	0.064	0.079	0.074	0.064	0.064	0.068	0.063	0.069

The data of this present experiment may serve to throw some light on this question. The results and tabulations were derived

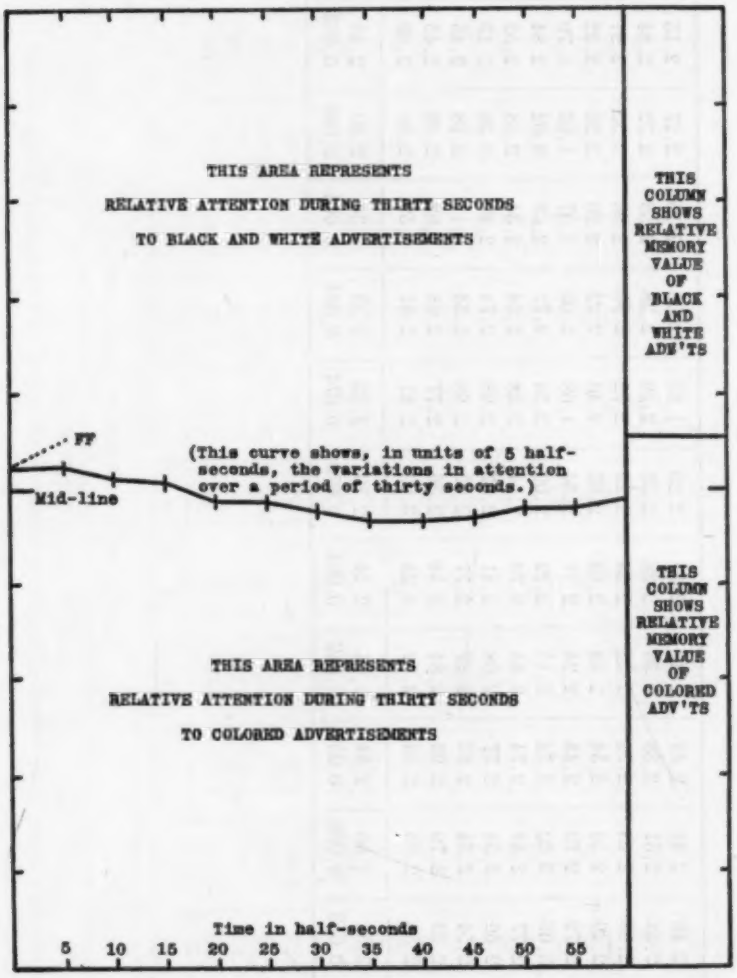

THIS AREA REPRESENTS
RELATIVE ATTENTION DURING THIRTY SECONDS
TO BLACK AND WHITE ADVERTISEMENTS
THIS COLUMN SHOWS RELATIVE MEMORY VALUE OF BLACK AND WHITE ADV'TS
FF
(This curve shows, in units of 5 half-seconds, the variations in attention over a period of thirty seconds.)
Mid-line
THIS COLUMN SHOWS RELATIVE MEMORY VALUE OF COLORED ADV'TS
THIS AREA REPRESENTS
RELATIVE ATTENTION DURING THIRTY SECONDS
TO COLORED ADVERTISEMENTS
Time in half-seconds
5 10 15 20 25 30 35 40 45 50 55

FIG. 2

in exactly the same way as in the case of pictures of people and with the same subjects. They are presented in table 3 and figure 2.

From a consideration of the table and the figure it is apparent that color is surprisingly inferior in attention power. Even in attracting initial attention (FF in the figure), for which it is ordinarily used, it is not nearly as effective as pictures of people. What initial advantage it does possess it loses very rapidly, so that after ten seconds the black and white competitor comes to be slightly superior, as is indicated by the fall of the curve for color below the mid-line on figure 2. Color does, indeed, show considerable memory value, about that found by Franken and Hotchkiss, but in general the use of color as an attention device under modern advertising conditions would hardly seem to be warranted when one considers the great extra cost. The truth is, perhaps, that not color as such but difference or novelty or change are the basic attention factors and in a magazine where colored advertisements prevail any one advertisement need not expect to profit especially in the way of extra attention. This does not invalidate its use for artistic purposes, of course, and from this point of view color may well be worth all it costs. It is likewise probable that in media where the use of color is infrequent its utility may be greater, as may also be the case where it is employed in some unusual or startling manner. In general we would merely urge here that it should not be too blindly accepted as a potent attention device.

CONCLUSIONS

1. Pictures of people are effective attention devices in advertisements. Interest in them increases for a considerable time and they have a high memory value.

2. The results of this experiment show color not to be so effective as an attention device as is generally supposed. It neither attracts great initial attention nor serves to hold interest. Whatever its general utility, in the absence of further experimental evidence it should not be regarded as an especially valuable attention factor under modern advertising conditions.

NOTES ON RATING

PERCIVAL M. SYMONDS

Teachers College, Columbia University

RATING VERSUS RANKING

More and more use is being made of rating scales in personnel work in industry and in education for the quantitative determination of otherwise unmeasurable traits and qualities. Although rating scales are generally preferred, ranking also is sometimes used and has generally been considered the more reliable. By rating I mean the assigning of a position on a scale of already defined steps. Rating is the method used in assigning a value to a composition on the Hillegas Scale or to a sample of handwriting on the Thorndike Scale. Rating is the method we use in the assertion that a child is in excellent, good, fair or poor health, the adjectives used roughly defining the steps on a four-point scale. By ranking I mean the placing of individuals in order as in arranging a gymnasium class in order of height or in sorting the cards in a pack in order from ace to king. If we think of individuals as differing in qualities in certain ways as from little to much, good to bad, desirable to undesirable, we may rank persons in such qualities as leadership, cleanliness, sportsmanship, and the like. At one end we place the individual with little of the trait in mind and at the other end the individual with much of the trait, trying to make each individual in the series exhibit more of the trait than his predecessor.

It is commonly thought that ranking is more reliable than rating because it demands a closer comparison. Allport (Social Psychology, 1924, p. 128) says "The ranking method, available only in rating an actual group all of whom are known to each rater, is *more accurate*. . . . . Comparisons are

more concrete and definite than in the scoring method." In rating, comparison is made with roughly defined steps or classes in the scale and the individual is placed in the nearest fitting class according to the opinion of the rater. But in ranking, one individual is compared directly with another.

The following experimentation was made to determine the facts. Two seventh grade teachers rated on a graphic scale a class of forty pupils on seven traits and seven habits. The school was organized on the departmental system so that the class was seen by the two teachers equally long and equally often. After a week the same teachers ranked the same pupil on the same traits and habits. An exactly similar program was carried out by two other teachers in the same school except that the ranking was done first and the rating second to allow for practice effect in our conclusions if there should be any. The traits and habits used were:

Honesty	Copies another's work
Obedience	Conforms to rules
Courtesy	Avoids passing in front of others
Orderliness	Keeps desk in good order
Cleanliness	Has clean hands
Sportsmanship	Is a good loser in a contest
Promptness	Puts away work quickly

The traits and habits were chosen primarily because of their prominence in school life and their availability for observation rather than for their importance.

The test of the relative reliability of the ratings as against the rankings was to compare the coefficients of correlation of the two teachers working on the same classes on any trait. That is, teacher A correlated with teacher B in *rating* honesty 0.472. The same two teachers correlated in *ranking* the same group in honesty 0.396. The average of twenty-eight coefficients by rating was 0.438 and the average of twenty-eight coefficients by ranking 0.445 (using Pearson's correction to Spearman's ρ which is obtained in using the Spearman rank formula for correlation). The two methods give nearly identical results. In another similar set of ratings and rankings

found in another school the same result held—rating and ranking were equally reliable. Also in a college class experiment in rating and ranking the difficulty of words in spelling, ranking had an average correlation with the true difficulties on the Ayres' spelling scale of 0.8966, rating had an average correlation with the true difficulties of 0.8973, and a second ranking a correlation of 0.8842. We may conclude then that rating and ranking are equally reliable.

Knight (1924) says,

> Judges who are presumably equally competent rate the same person as almost wholly honest, half honest, rather dishonest, or an out-and-out cheat. This disagreement is the product of legitimate variations in knowledge about the person and of varying standards of honesty held by the judges themselves. The way to avoid this pitfall is not to use a rating involving concepts of virtues, but a system of ranking persons according to their relative merits. The question should be not, "How honest is he?" but "Is he more or less honest than this person or that?"

Knight's objection to rating scales is not wholly an objection to their differences in reliability—he is objecting to the tendency to rank high or low on a rating scale as a whole. One teacher might give a class an average rating of 4 on a rating scale in honesty and another teacher give the same class a rating of 2 and yet the two might correlate perfectly. In using the ranking method we can transmute the ranks into sigma indices by means of the tables such as are found in Chu (1922) or into percentiles by means of the tables such as are found in Haught (1923), and can assume that the means of different groups are the same, an assumption which probably is nearer the truth in most cases than taking the means to be as different as may often be found on rating scales. On the other hand, every one that I have asked tells me that rating is more pleasant than ranking. The graphic rating scale seems to be most easy to use. Ranking when the groups become large grows confusing and the effort to make fine distinctions between individuals becomes irksome. Our results show that the fine distinctions simply are not made.

I should like therefore to suggest an improvement in rating schemes which would keep the advantages of both methods. Most rating schemes are like this:

RATING CHART A

Pupil's name.....Date.....
School.....Grade.....
Rated by.....

Directions for using the Rating Chart

1. Let these ratings represent your own judgment. Do not confer with anyone in making them.
2. In each trait or characteristic named below compare this pupil with the average pupil of the same age.
3. In any particular trait disregard every other trait except that one. Many ratings are rendered valueless because the rater allows himself to be influenced by a generally favorable or unfavorable impression which he has formed of the person rated. Do not rate a pupil high on all traits simply because he is exceptional in some. Children are often very high in some traits and low in others.
4. Place a cross somewhere on the line running from "very high" to "very low" to indicate this child's standing in each quality. *You may place your cross at any point on the line.* It is not necessary to locate it at any of the division points or above any descriptive phrase.
5. Do not study too long over any one trait. Give for each the best judgment you can and go on to the next.
6. Give a rating for every trait.
7. The ratings will be held strictly confidential.

HEALTH—Is he generally healthy and vigorous?

Bad	Poor	Average	Good	Excellent

LEADERSHIP—Does he take the lead in school affairs or does he follow others?

Always follows others	Rather tends to follow	Average	Rather tends to be a leader	Masterly, not easily influenced

The usual method with such a scale is to go through the traits for one individual and then turn the sheet to the next individual. Even though instructed to rate all individuals on one

trait before turning to the next trait, no direct comparisons can be made since only one page is visible at a time. To make use of the close comparison which is demanded in ranking the following rating scheme is recommended:

RATING CHART B

Rated by.....Date.....
School.....Grade.....

Directions for using the Rating Chart

1. Let these ratings represent your own judgment. Do not confer with anyone in making them.
2. In each trait or characteristic named below compare this pupil with the average pupil of the same age.
3. In this particular trait disregard every other trait except this one. Many ratings are rendered valueless because the rater allows himself to be influenced by a generally favorable or unfavorable impression which he has formed of the person rated. Do not rate a pupil high on this trait simply because he is exceptional in some. Children are often very high in some traits and low in others.
4. Place a cross in one of the compartments running from "very high" to "very low" to indicate each child's standing in this quality.
5. Do not study too long over any one child. Give for each child the best judgment you can and go on to the next.
6. Give a rating for each child.
7. The ratings will be held strictly confidential.
8. Try to let the percentages guide you as to the number of crosses to fill in each compartment.

Trait: HEALTH

Is he generally healthy or vigorous?

PUPIL	4% VERY BAD	11% BAD	21% POOR	28% AVERAGE	21% GOOD	11% VERY GOOD	4% EXCEL-LENT
Charles.....							
William.....							
George.....							

Such a rating chart keeps the desirable features of the graphic rating chart and also permits the close comparison of ranking.

To be sure the ratings for any one individual are spread over several sheets, but as in any group test the ratings for the several tests for any individual may easily be brought together when the numerical values of the graphic ratings are measured.

Whatever such a chart would contribute toward increased rating accuracy, it would probably in no way do away with the necessity of having three or more qualified and competent persons rate independently in order to have a final pooled rating whose reliability could be depended upon for individual diagnosis.

THE SIZE OF THE HALO EFFECT IN RATINGS

For some time users of rating devices have noticed that ratings of human traits are inaccurate because the rater's judgment of the trait in a certain individual is influenced by his general impression of that individual. So far as I am able to determine Webb in his excellent and only too little read monograph "Character and Intelligence" was the first to notice this halo effect in 1915. On page 70, he says "Let us suppose, for instance, that the observers, in estimating the intelligence qualities, are biased in the direction of marking subjects who possess other desirable qualities too highly, and vice versa." Thorndike (1920) seems to be the first to notice this phenomenon in this country and to use the term halo in identifying it. At that time he said, "The magnitude of the constant error of the halo, as we have called it, also seems surprisingly large, though we lack objective criteria by which to determine its exact size." In this paper partial correlations are used to determine the size of this halo effect.

As previously described, correlations were found between the ratings of the two teachers on any trait or habit. That is, teacher A correlated with teacher B in the ratings on honesty 0.472. Composite ratings of the traits were found by adding together the ratings of the seven traits for each child. This composite was taken as the general impression of the child by the rater. Then partial correlation coefficients of the second order were found for the ratings on each trait with the composites for each of the raters made constant.

1. The halo effect which heretofore has been merely assumed is here demonstrated to be a reality. Seven trait ratings have their correlations raised by 0.245 by the halo effect. Seven habit ratings have their correlations raised by 0.177 by the halo effect.

TABLE 1

Coefficients of correlation between ratings of teacher A and teacher B

	ZERO ORDER COEFFICIENTS	COEFFICIENTS WITH COMPOSITES OF A AND B MADE CONSTANT	DIFFERENCE
Honesty	0.471	0.193	−0.278
Obedience	0.387	−0.036	−0.423
Courtesy	0.406	0.110	−0.296
Orderliness	0.190	0.104	−0.086
Cleanliness	0.470	0.553	0.083
Sportmanship	0.360	0.000	−0.360
Promptness	0.448	0.094	−0.354
Average	0.390	0.145	−0.245
Copies another's work	0.587	0.470	−0.117
Conforms to rules	0.313	0.052	−0.269
Avoids passing in front of others	0.306	0.116	−0.190
Keeps desk in good order	0.331	0.110	−0.221
Has clean hands	0.338	0.277	−0.061
Is a good loser in a contest	0.488	0.264	−0.224
Puts away work quickly	0.256	0.101	−0.155
Average	0.377	0.200	−0.177

2. With our data we are not able to demonstrate with statistical sureness that there is a real difference in the rating of traits and habits. But there do seem to be differences which are real. Obedience loses 0.423 with halo made constant while cleanliness gains 0.083 with halo partialled out. The difference is not one of habits versus traits. Possible reasons for a large halo effect in any trait or habit are (1) the trait or habit is one which is not easily observed, (2) the trait or habit is one which is not commonly observed or thought about, such

as one which is not usually emphasized in a classroom, (3) the trait or habit is not clearly defined, (4) the trait or habit is one which involves reactions with other people rather than mere personal behavior, (5) the trait or habit is one with high moral importance in its usual connotation.

The ratings mentioned in this paper were made by four able teachers of Kaahumanu School, Honolulu, Hawaii: Miss Grenville Hatch, Mrs. E. Lemon, Miss Ilda Astleford and Miss E. Astleford.

REFERENCES

ALLPORT, F. H. 1924 Social Psychology.

CHU, J. P. 1922 Chinese students in America: Qualities associated with their success. Teachers College Contributions to Education, no. 127.

FREYD, M. 1923 The graphic rating scale. Jour. Educational Psychology, vol. 14–2, February, 1923, pp. 83–102.

HAUGHT, F. B. 1923 A scheme for combining incomplete rankings. Jour. Applied Psychology, VII–2, June, 1923, p. 172.

KNIGHT, F. B. 1924 Analysis of teaching and teachers. Jour. Educational Research, vol. X–5, October, 1924, p. 224.

PATERSON, D. G. 1923 Methods of rating human qualities in psychology and business. Annals American Academy of Political and Social Science, vol. CX–199, November, 1923, pp. 81–93.

THORNDIKE, E. L. 1920 A constant error in psychological rating. Jour. Applied Psychology, 4: pp. 25–29.

WEBB, E. 1915 Character and intelligence. British Jour. Psychology, no. 3, 1915, pp. 1–99.

A TEST FOR CLASSIFICATION OF STUDENTS IN CHEMISTRY

WINIFRED F. HYDE
AND
J. P. GUILFORD
University of Nebraska

The test here discussed gave such gratifying results that the writers feel justified in reporting their experience with it. The test is the outgrowth of an attempt to solve a problem which was brought to the department of psychology by an instructor in chemistry.

At the above-mentioned university, there are two beginning courses in chemistry: Chemistry 1—five hours credit—a course for those who have had no high school chemistry; and Chemistry 3—three hours credit—for those who have had one year of high school chemistry; both courses to prepare the students for the same advanced work. Chemistry 3 has always had an undue number of failures, many of whom for some reason or other needed more thorough drilling in the fundamentals of the subject, and had they entered Chemistry 1, would have stood a better chance of making a passing mark. According to custom, at mid-semester those who showed need of more elementary work in the subject or a prospect of failure were advised to go into Chemistry 1, but with the inconvenience, in some cases, of changing their semester schedule. The instructor wanted to know if a test could be devised which would detect these students so that they might be advised to enter Chemistry 1 at the beginning of the semester.

Two main factors are probably responsible for failure in Chemistry 3, namely, lack of ability for chemistry, and lack of preparation in high school. Our test, accordingly, contained two parts, Part I, which we will call a Psychological Test, and

Part II, a Chemistry Information Test. The instructor had found through experience that the students are most likely to fail in the solution of chemical problems. These problems are of a rather uniform type, involving the use of ratio and proportion, and requiring the ability to see analogous relationships. Our psychological examination therefore emphasized that type of problem and was composed of four tests:

Test 1, consisting of twenty arithmetic problems with the insertion of a few simple chemical problems which were typical of the kind the student would have to solve in the course.

Test 2, a word analogies test, a very few of them involving chemical terms.

Test 3, a form analogies test, such as is used in the Thorndike test for college freshmen. This presents the same kind of problem as in Test 2, except in the form of non-verbal material.

Test 4, a number series completion test, such as is used in the Army Alpha and other group tests.

The Chemistry Information test was composed of the following:

Test 5, a sentence completion test, the blanks to be filled with chemical terms found in a miscellaneous list at the top of the page.

Test 6, a true-false test, containing twenty statements of well-known facts in chemistry.

The total possible score in the psychological test is 70, weighting the arithmetic test by doubling the number right, and for the chemistry information test, 55 points, or a total combined score of 125. Since the aim was to pick out the least capable chemistry students, the tests were made very simple and the time long enough so that only the slowest could not finish. The total time for work on the tests amounted to about twenty-two minutes.

Two alternative forms were provided, and although an attempt was made to make them of equal difficulty, Form I proved to be a little more difficult in the information part, as will be seen in the table of medians (table 1).

The range was sufficiently great to allow for good discrimination of ability as shown in the distribution curves in figure 1.

The curve for the psychological test is skewed to the right, which makes for better discrimination among those in the lower quartile,—a condition we tried to produce by making the test easy. But the curve for the chemistry information test is skewed in the opposite direction, which makes it more easy to pick out the very superior rather than the very inferior in this test. The test is too difficult for the purpose for which it was intended.

TABLE 1

Medians

	PSYCHOLOGICAL	CHEMISTRY	COMBINED	NUMBER OF CASES
Form I	41.62	14.83	56.33	107
Form II	41.55	18.8	60.55	106

But the proof of the value of a test is the pragmatic one: does it work? The correlation of the tests with the semester grades is very significant. It is disturbed by only one fact, that seventeen of the students at the beginning of the semester

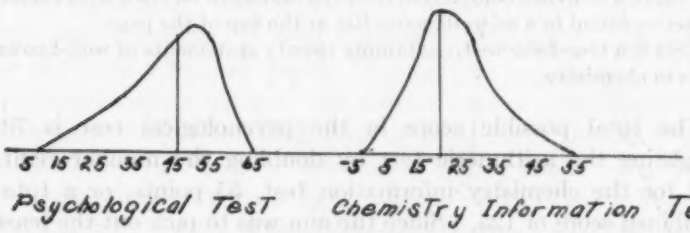


FIG. 1

went into Chemistry 1 upon the advice of the instructor. The latter was given the names of the lowest ten per cent in the two tests, and it was upon this basis and a conference between student and instructor that they were advised to take the more elementary course. But neither the instructor nor any of those who had anything to do with grading the students in Chemistry 3 knew any of the scores of the students until the end of the semester.

We have included in our correlations eight of the students who went into Chemistry 1 and failed, assuming that they would have failed also in the more difficult course, but did not include the nine others who passed in Chemistry 1 because we are not sure that they would have passed in Chemistry 3. The coefficients of correlation are given in table 2.

The correlations would tend to show that the psychological test is a much better criterion of ability to succeed in general chemistry than is the test for chemical information retained

TABLE 2

TEST		CORRELATION WITH GRADE	P.E.
Psychological	Form I	+0.448	±0.0537
	Form II	+0.649	±0.0388
	Both	+0.566	±0.0455
Chemistry Information	Form I	+0.204	±0.0648
	Form II	+0.437	±0.0545
	Both	+0.331	±0.0599
Combined score	Form I	+0.453	±0.0535
	Form II	+0.583	±0.0435
	Both	+0.517	±0.0493

from a high school course. This agrees well with the opinion of the instructor, that failure in chemistry is not so much due to lack of preparation in high school as it is to lack of native ability. It also accounts for the fact that eight of the seventeen who made low scores in the test failed although they went into the easier course.

But coefficients of correlation do not always tell us just how well we can select the potential failure, or what mistakes we might make by selecting an arbitrary passing limit for a test. Since the psychological test correlates better with success in chemistry than does the chemistry information test, and even better than the two combined, it would be wise to take the score in that test alone as a criterion of success in the course. Table 3 shows what per cent of the students were failed or

conditioned in Chemistry 3 for each degree of success in the psychological test.

Thus, those who fall below 20 in the test are almost sure to fail, and also half of those between 20 and 30. It would be

TABLE 3

SCORE	CASES	FAILURE	CONDITION	PER CENT F AND C
				per cent
10–19	3	2	1	100
20–29	18	7	2	50
30–39	48	6	4	21
40–49	76	3	8	14
50–59	37	0	1	3

TABLE 4

STUDENT	SCORE			INSTRUCTOR'S STATEMENT
	Psychology	Chemistry	Total	
C N. T.	47	19	66	Not overly bright
C J. B.	55	26	81	Lazy
C J. M.	44	20	64	Does not work
C J. M.	47	18	65	Lazy. Poor record. Pals with boys who do not work
C R. M.	49	14	63	Really should not have been conditioned. Has worked it off with good examination. Football. Doing well the second semester
C L. M.	44	22	66	No statement
C R. G.	42	24	66	Had been out of school. Hard to adjust himself
C P. B.	47	13	60	"Sluffer." Bright
C A. S.	42	19	61	Outside work handicaps
F R. G.	41	25	65	All right but sickly

well, then, to advise all those below 30 to take the more elementary course unless they showed sufficient reasons on other grounds why they are able to pass successfully the more difficult course.

We cannot overlook the fact, however, that this method would eliminate only a third of the failures, as can be seen from the table. Of the 22 who made a score above 30 and were failed or conditioned, many reasons other than lack of ability conspired to prevent them from coming up to our expectations. Those who fall between 20 and 30 may be classed as of doubtful ability. But the twelve who made a score above 40 (the median is 42) are surely not lacking in ability. We asked the instructor for statements with regard to these twelve students, and invariably his opinion agreed with our supposition. The statements are given in table 4.

TABLE 5

STUDENT	SCORE			INSTRUCTOR'S STATEMENT
	Psychology	Chemistry	Total	
J. D.	16	4	20	Hardworking, slow
C. D.	24	−5	19	Older student; slow; hard worker
J. K.	18	14	32	Very slow
B. E.	26	11	37	Seems capable; much interested in chemistry

Four students with scores below 30 made a grade of 80 in Chemistry 1. A statement from their instructor (table 5) helps to clear up these cases.

Three of these, because they were slow, were evidently handicapped in the tests, which although allowed a generous amount of time, hurried them too much to be able to do their best work.

Taking all these things into consideration, the tests proved their value to classify students in the two chemistry courses. The question is raised whether, since it is native ability and not preparation in high school which is most essential for success in chemistry, some of the students of very high ability who have had no high school chemistry could not carry the more difficult and briefer course and thus save time in their university work.

Such a plan has not yet been tried, but our results would seem to justify such an experiment. The results also verify the fact that purely psychological or intelligence tests are on the whole a better criterion of success in college than are entrance examinations upon material learned in high school.

CONCLUSIONS

1. By means of a psychological test given at the beginning of the school year and taking not over twenty minutes of time, we are able to eliminate two-thirds of the failures in Chemistry 3, who lack the ability, and to save about half of these from failing in beginning chemistry by placing them in a more elementary course.

2. A psychological test given to freshmen is of more value for predicting success in college than are examinations on high school subjects.

NOTES AND NEWS

FELLOWSHIPS IN THE PSYCHOLOGY AND EDUCATION OF THE PRE-SCHOOL CHILD

The Merrill-Palmer School of Detroit, Michigan, is offering for the year 1925–1926, three fellowships for graduate study in the field of the psychology and education of the pre-school child.

A preference will be given to students who are candidates for higher degrees. Some universities, by special arrangement, are willing to accept work done at this school as part of the requirements for a higher degree. It is entirely possible to do experimental work which will serve as the basis of a thesis for a degree. The fellowships pay $1000 each.

The Merrill-Palmer School offers the following facilities:

Two nursery school centers: One for children between eighteen months and three and one-half years of age, with an enrollment of twenty-five, and the other for children between three and five years, with an enrollment of thirty-five.

A psychological laboratory in which all of the children of both schools are carefully measured and re-measured. This laboratory has now completed the standardization of a new series of tests for children between eighteen months and five and one-half years, by half-year periods. The standardization was based upon between five and six hundred children.

A consultation center to which physicians, parents, and social workers refer young children for diagnosis and advice. The center is equipped to make mental measurements, to give complete physical examinations, if these are not already provided for, and to make social investigations. Problems of behavior are analyzed, and advice given about treatment. Parents are advised about the educational environment of young children.

The school is equipped to offer excellent courses in the health and nutrition of young childhood, which may be taken as a minor by students specializing in psychology.

Applications giving full particulars of the academic background and of previous professional experience, if the candidate has had experience in schools, clinics, or social agencies, should be mailed to the Merrill-Palmer School before the end of the present academic year.

HELEN T. WOOLLEY, PH.D.,
Psychologist and Director of the Nursery School.

Dr. Donald A. Laird, Colgate University, will be the chief psychological editor of *Industrial Psychology*, a new scientific publication which will make its appearance in January, 1926. Mr. Percy S. Brown, works manager of the Corona Typewriter Company and President of the Taylor Society will be chief industrial editor. Each issue will contain 80 pages of reading matter devoted to the problems of employment, personnel, industrial relations and other problems connected with industry. Hopeful but impossible notions are to be culled by having both editors pass on each paper before publication. A special rate of $4.00 a year is offered to those subscribing before the first issue. Address *Industrial Psychology*, Box 34, Hamilton, N. Y.

The JOURNAL OF APPLIED PSYCHOLOGY extends a cordial welcome to this new contribution.

The Eighth International Congress of Psychology will be held at Groningen, Holland, from September 6 to 11, 1926. Invitations have been sent to about 200 well-known psychologists. Others who desire to attend are requested to apply to the First Secretary, F. Roels, 86 Maliebaan, Utrecht, Holland.

On May 14 to 16, 1925, at Ann Arbor, was held a meeting of the Taylor Society, an international society to promote the science and the art of administration and of management, under the auspices of the University of Michigan. Of special interest was a paper presented by C. S. Yoakum, Professor of Personnel Management, University of Michigan, on Experimental Psychology in Personnel Problems. This was followd by discussion led by Prof. A. W. Kornhauser, School of Commerce and Administration, University of Chicago, and Lillian M. Gilbreth, Frank B. Gilbreth, Inc., Monclair, N. J.

Dr. Harry D. Kitson has accepted a position in the Department of Psychology at Columbia University. He has been connected for some years with the Department of Psychology, University of Indiana, Bloomington, Ind.

Dr. Ada Hart Arlitt, formerly of Bryn Mawr College, now with the Cincinnati Central Clinic, will become head of a Department of Child Care and Training to be established next year in the University of Cincinnati.

BOOK REVIEWS

G. M. Ruch. *Improvement of the Written Examination.* Scott, Foresman and Co. x + 193 pp. Price $1.80.

During the past two decades there has been a growing amount of experimental data on the reliability—or rather the unreliability—of teachers' marks. The general conclusion common to all the studies reported in the technical literature is unaminous—examination grades and semester grades vary widely not only from one instructor to another, but the grading of any instructor is highly unreliable. In other words, the measure of relative school success or failure most commonly used is in point of fact so largely a chance affair that the proper placement of a given individual pupil is bound to be grossly in error in a large proportion of the cases. The measurement of the achievement of educational goals is by the same token subject to the same limitations insofar as these measurements are a function of the typical examination grade.

All of this is a matter of a b c's for the educational technician. Yet the fact remains that very little has been offered by way of remedy. Wood, Monroe, McCall, Toops, Knight are a few of the names that come to mind relative to those who have given some constructive suggestions. What has been needed is a book which would summarize the results of the necessarily piecemeal activities of the investigator, evaluate them critically, analyze the aims of education relative to examinations, and map out a program about which the garden variety of teacher could do something. Standardized tests, useful as they proved and will continue to be, do not answer the need. Apart from other weighty considerations, they are not adapted to the relatively short time intervals at which successive measurements are desirable.

The book here under review performs a notable service in line with the above considerations. Dr. Ruch has not only admirably summarized heretofore scattered conclusions, but adds a considerable amount of experimental data on his own account to show the relative merits of different types of examinations. The theoretical and technical discussion is well within the grasp of the average classroom teacher and is marked by scientific conservatism rather than any tendency to enthusiastic overstatement. There is no attempt to buttress absence of experimental data with dogmatism. The educational practioner in the classroom will be most grateful for the inclusion of directions for preparing adequate types of examinations together with much illustrative material in various kinds of subject matter such as chemistry, biology, literature

and composition, French, Latin, physics, spelling, reading, history, and manual training.

The six chapter headings will indicate the content and organization of the book: Chapter I, Functions of Written Examinations; Chapter II, The Criteria of a Good Examination; Chapter III, Sources of Error in Written Examinations; Chapter IV, Types and Construction of the Newer Objective Examinations; Chapter V, Experimental Studies of Several Types of Objective Examinations; Chapter VI, Statistical Considerations Related to Examination Technique. The Appendix contains a complete high school content examination as well as some of the illustrative material referred to above. The summaries and questions at the close of each Chapter are exceptionally well worked out.

In the judgment of the present reviewer the book should find a place in all elementary courses in educational tests and measurements as well as in teachers' reading circles, study groups, etc. To quote from the Preface by Dr. Ernest Horn: "There has now developed a new type of test program. . . . . These new and more perfect methods of testing fill a need which cannot be met by the use of standardized tests. . . . . Dr. Ruch has performed a real service in presenting the evidence concerning the reliability of various forms of tests and in illustrating so clearly the different ways in which these tests may be prepared. There is no question . . . . but that this book marks out the next big field of improvement in testing the results of teaching."

H. H. Remmers,
Purdue University.

John Monk Saunders and George Palmer. *Brain Tests.* G. P. Putnam's Sons: New York, 1925. Pp. 95. $1.50.

Adieu to crosswords! Here we find a small book, with pencil attached, by which we can test our brains and ascertain just how our mind works. Part of the jacket resembles the crossword cover but the book itself is filled with abbreviated psychological tests which could not have been made more intriguing.

A key and an "average" score is given for each test in the back of the book whereby the reader may check up his powers of observation, range of information, ability to follow directions and tread through a maze, make change or try his skill at geometric figures.

I predict considerable success for this book, and think that psychology will receive a little more popularization from it. What it will tell the reader about himself is difficult to say, but he will certainly find it interesting.

Donald A. Laird, Ph.D.,
Colgate Psychological Laboratory.

NATALIE KNEELAND. *Cases in Retail Salesmanship.* Merchandise Manuals, A. W. Shaw Company, 1924. Pp. 189. Price $1.50.

Any one who can imagine the ignorance and trepidation with which a young salesgirl approaches her first customer in a department store can understand how great is the need for aids for retails salespersons. A number of aids are already available. There are several textbooks on retail selling besides the long list of books covering salesmanship in general. Courses are offered by correspondence, and training is offered by most department stores. In addition to these general helps for salespersons there are a number of books giving technical information about the goods sold in department stores. In spite of the volume of such aids there is still room for another kind of help. Salespersons find it hard to pass general principles of selling to concrete instances. And they find the masses of technical information about goods quite beyond their power to assimilate. Recognizing these difficulties, Doctor Charters and his associates in the Research Bureau for Retail Training, of Pittsburgh have prepared a series of sales manuals that have as their aim the presentation of facts in their concrete setting and the recommendations of practises in the concrete.

The number of manuals announced by the publisher to be in preparation is seventeen in all. The first one, "Cases in Retail Salesmanship," is an attempt to apply the case method of instruction to all the departments of a department store. It is a collection of 135 accounts of actual sales, some reported by students of the Bureau, some by shoppers. They are reported in dialog form. The sales were selected as illustrative of certain types of situations such as "Knowing what is in stock;" "Selecting merchandise appropriate to the customer's taste;" "Knowing when to approach the customer;" "Handling two customers at once." For example,under the heading, "Finding out what the customer wants by asking questions," appears the following dialog:

Salesperson—Is there something for you, madam?
Customer—Well, I was just trying to decide.
Salesperson—Are you looking for yourself?
Customer—No, I was thinking about a dress for my daughter.
Salesperson—Does she require a girl's or junior size?
Customer—She is just my size.
Salesperson—We have some lovely new sport silk frocks. Would you care for something like this?
Customer—Oh, that's just what she wants. I heard her speak of having one, the other day.
Salesperson—Is she dark or light?
Customer—Her hair is red.
Salesperson—Oh, then this same dress in tan and brown would be lovely on her. Isn't that a lovely color? (showing dress).

Customer—And how pretty she would look in it! She loves tan and brown so much.

Salesperson—The girls are wearing these dresses so much now for sport wear.

Customer—I believe I'll let you send that up.

At the close of each account questions are presented which are intended to lead the student to analyze the sale and determine its strong and weak points. References are also given to several books on retail selling.

Several errors crept into this manual: On page 30 "fallacy" is misspelled; on page 33, yoke; on page 52 the question, "When will this be out?" should be attributed to the customer instead of to the salesperson; on page 178 appears the phrase "principles . . . . was;" on page 183, "iridescent" is misspelled.

NATALIE KNEELAND. *Waists.* Merchandise Manuals. A. W. Shaw Company, 1924. Pp. 126. Price $1.50. *Hosiery, Knit Underwear and Gloves.* Pp. 126. Price $1.50.

FREDONIA JANE RINGO. *Girls' and Juniors' Ready-to-Wear.* Merchandise Manuals. A. W. Shaw Company, 1924. Pp. 160. Price $1.50.

The manuals devoted to various commodities are compilations of the facts about the commodities which a salesperson would find most useful in "talking up" her goods. For pedagogical reasons they are arranged in the form of questions and answers, such as:

Question: "What different qualities of cotton are used in underwear?"

Answer: "There are many grades of qualities of cotton. The most important to know are the Sea Island, Egyptian, American Upland, and Peruvian."

Then follow paragraphs on each of these. Each section is ended by "Test Questions." A section in each manual is devoted to Methods of Teaching, much of the material in this section is the same in all the manuals.

A unique and much-needed section of the manual on Girls' and Juniors' Ready-to-wear contains a vocabulary for describing values. A "black-list" shows words that salespersons should avoid, such as cheap, classy, clever, cute, elegant, grand, sweet, etc. Instructions are given to use words conveying definite ideas instead of general ideas. For example, instead of saying, "This is a lovely blue," one should say, "This is Delft blue, or cadet blue." A long list of words that may be used in describing texture includes airy, clinging, filmy, fleecy, sheer, supple, etc. Words describing different styles are artistic, conservative, draped, fitted, flaring, jaunty, ornate, severe, swagger, ultra, etc. For

the sake of this section alone every retail salesperson should study these books. The practical value of all the aids offered by this series of manuals will undoubtedly insure their wide acceptance by those who are in charge of training retail salespersons.

HARRY D. KITSON,
Indiana University.

Helmholtz's Treatise on Physiological Optics. Translated from the Third German Edition. Edited by James P. C. Southall. Vol. I. The Optical Society of America, 1924. Pp. xxi, 482.

At a meeting of the Optical Society of America in 1921 it was proposed to commemorate the hundredth anniversary of the birth of Helmholtz "in a more useful and substantial way by bringing out the long-delayed English translation of the Physiological Optics, and accordingly a special committee was appointed to have charge of this business." Let it be said at once that the Optical Society, and more especially Professor Southall and his collaborators in the arduous task of translation and editing, have performed a notable and valuable service by publishing in English "one of the four or five greatest monuments of human genius in the scientific line."

The French have had their translations of Helmholtz sooner than we have. The *Handbuch der physiologischen Optik* first made its appearance, in four installments, between the years 1856 and 1866. Within a year it was published in French by E. Javal and N. Th. Klein. *Die Lehre von den Tonempfindungen* appeared in 1862 and was translated by G. Guéroult six years later and supplemented by an appendix after the third German edition in 1874. But in biding out time we have come into possession of much more serviceable versions of Helmholtz's works. The present edition of the *Optik* by Professor Southall, even if for no other reason than that it is a translation of the third German edition, is more valuable than the French one. And A. J. Ellis' version of the *Tonempfindungen* is in a class by itself.

Professor Southall and his committee decided that the third German edition was beyond question the one to be reproduced in English. This edition was brought out in three volumes after Helmholtz's death, under the editorship of W. Nagel, in collaboration with Professors A. Gullstrand and J. von Kries, between 1909 and 1911. Nagel followed the text of the first edition instead of the second, which meant the exclusion of certain of König's researches but gave, on the other hand, a more faithful picture of the extraordinary brilliance with which Helmholtz dealt with phenomena which took him into the fields of physics, mathematics, physiology, psychology, and philosophy. The edition was expanded to nearly double the size of the original by the addition of new material in the field of physiological optics contributed by the editors.

Readers of the English translation, then, will have all of this material before them, and in addition, significant new contributions which have been made to the subject in the last decade and numerous excellent foot-notes by Professor Southall and his collaborators.

The first volume, Anatomical Descriptions and Dioptrics of the Eye, contains, in addition to all the material of the German edition, an entirely new chapter by Professor Gullstrand on his own remarkable researches in ophthalmoscopy; a bibliography of more recent literature on anatomy of the eye compiled by D. H. Hooker which, "although it does not pretend to be a complete list by any means," has nearly 200 titles; numerous new foot-notes and references to literature; and a portrait of Helmholtz which appeared in the third volume of the German edition. In the preparation of the first volume the editor was assisted in large measure by D. H. Hooker and L. D. Weld.

The English-reading scientific world is under a debt of gratitude to the American Optical Society for the appearance and very excellent preparation of Helmholtz's great work, and especially to Mr. Adolph Lomb without whose "continual advice and encouragement the achievement would never have been accomplished."

Helmholtz's Treatise on Physiological Optics. Translated from the Third German Edition. Edited by James P. C. Southall. Vol. II. Published by the Optical Society of America, 1924. Pp. x, 480.

The second volume of the third German edition of the *Optik* appeared in 1911, one year after the third volume. The delay was due to the illness of the editor, W. Nagel. After Nagel's death von Kries assumed the responsibility for the completion of the second volume. Nagel had already prepared parts of an appendix for the second volume on adaptation, twilight vision, and the duplicity theory, which were included by the new editor together with an appendix of his own on normal and anomalous color systems and theories of vision. These additions by Nagel and von Kries to the original text of Helmholtz are reproduced, of course, in the English translation of the second volume which now follows close upon the appearance of the first volume.

Psychologists will find the contents of the second volume, sensations of vision, of more immediate interest to them than those of the first volume. Here one has Helmholtz's great researches on visual stimulation, simple and compound colors and their attributes, contrast, etc., and the appendices by Nagel and von Kries. The English edition contains, in addition, a chapter by Dr. Christine Ladd-Franklin on the nature of color sensations. In this chapter Dr. Ladd-Franklin has discussed certain of the weak features of the Helmholtz theory of vision, reproduced the facts and diagrams derived from König's important researches, which are not elsewhere fully treated in the English trans-

lation since the editors decided to follow the third rather than the second German edition, and given an account of her own development theory of color sensations. The English volume also contains three entirely new notes especially prepared by von Kries in simultaneous contrast, the connections between various forms of color vision in dichromata, and methods for preparing the luminosity of light of different colors; and an impressive bibliography, compiled by the editor, of works relating to sensations of vision which have appeared since 1911.

What has already been said about the preparation of the first volume is equally true of the second. Professor Southall and his collaborators deserve the highest praise for bringing before English readers such an excellent version of Helmholtz's monumental work! In the preparation of the second volume the editor was assisted by H. Laurens, M. Dresbach, L. T. Troland, and E. J. Wall.

C. C. Pratt,
Harvard University.

NEW BOOKS AND PAMPHLETS RECEIVED[1]

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

Abnormal Psychology and Education. FRANK WATTS. D. Appleton and Company, 1924.

Challenge of Childhood, The. IRA S. WILE. Thomas Seltzer, 1925.

Coöperation in Adult Education. ELLEN C. LOMBARD. Home Education Circular No. 6. Bureau of Education, Washington, D. C. 34 pp.

Genetic Studies of Genius, Vol. I. LEWIS M. TERMAN. Stanford University Press, Stanford University, Calif. Price $5.00. 848 pp.

Graphic Methods in Education. J. HAROLD WILLIAMS. Houghton Mifflin Company, 1924. Price $2.00.

Human Conduct and the Law. MARY C. LOVE. George Banta Publishing Company, Menasha, Wisconsin. 318 pp.

Lefthandedness, A New Interpretation. LUIS PARSON BEAUFORT. The Macmillan Company, 1924.

Legal Provisions for Rural High Schools. WILLIAM R. HOOD. Bulletin No. 40, 1924, Bureau of Education, Washington, D. C. 60 pp.

Problems in Physical Education. JAMES FREDERICK ROGERS. Physical Education Series No. 5. Bureau of Education, Washington, D. C. 19 pp.

Psychotechnik der Beroepskeuze. J. L. PRAK. Bij. J. B. Wolters U. M., Groningen, Den Haag. 272 pp.

Purposive Speaking. ROBERT WEST. The Macmillan Company, 1924.

References on Education for Citizenship. Library Leaflet No. 30, January, 1925. Bureau of Education, Washington, D. C. 16 pp.

Statistical Method in Educational Measurement. ARTHUR S. OTIS. World Book Company, Yonkers, N. Y. 337 pp.

Statistics of State School Systems. FRANK M. PHILLIPS. Bulletin No. 31, 1924. Bureau of Education, Washington, D. C. 42 pp.

A Study of 260 School Consolidations. J. F. ABEL. Bulletin No. 32, 1924, Bureau of Education, Washington, D. C. 39 pp.

Ways of the Mind, The. Charles Scribner's Sons, 1925. 336 pp.

Yiddish Press, The. An Americanizing Agency. MORDECAI SOLTES. Teachers College, Columbia University. 1925.

---

[1] Mention here does not preclude further comment.

Studies from Columbia University, Teachers College

Changing Conceptions Relative to the Planning of Lessons. LOIS COFFEY MOSSMAN. No. 147, 1924.

Content and Form in Tests of Intelligence. EDWIN MAURICE BAILOR. 1924.

A Comparative Study of the Mental Capacity of Children of Foreign Parentage. MAY BERE. No. 154, 1924.

Detailed Factors in Latin Prognosis. ORLIE M. CLEM. No. 144, 1924.

Diagnostic Study of the Subject Matter of High School Chemistry, A. SAMUEL RALPH POWERS. No. 149, 1924.

Evaluation of English Literature in the High School. CHARLES SUMNER CROW. No. 141, 1924.

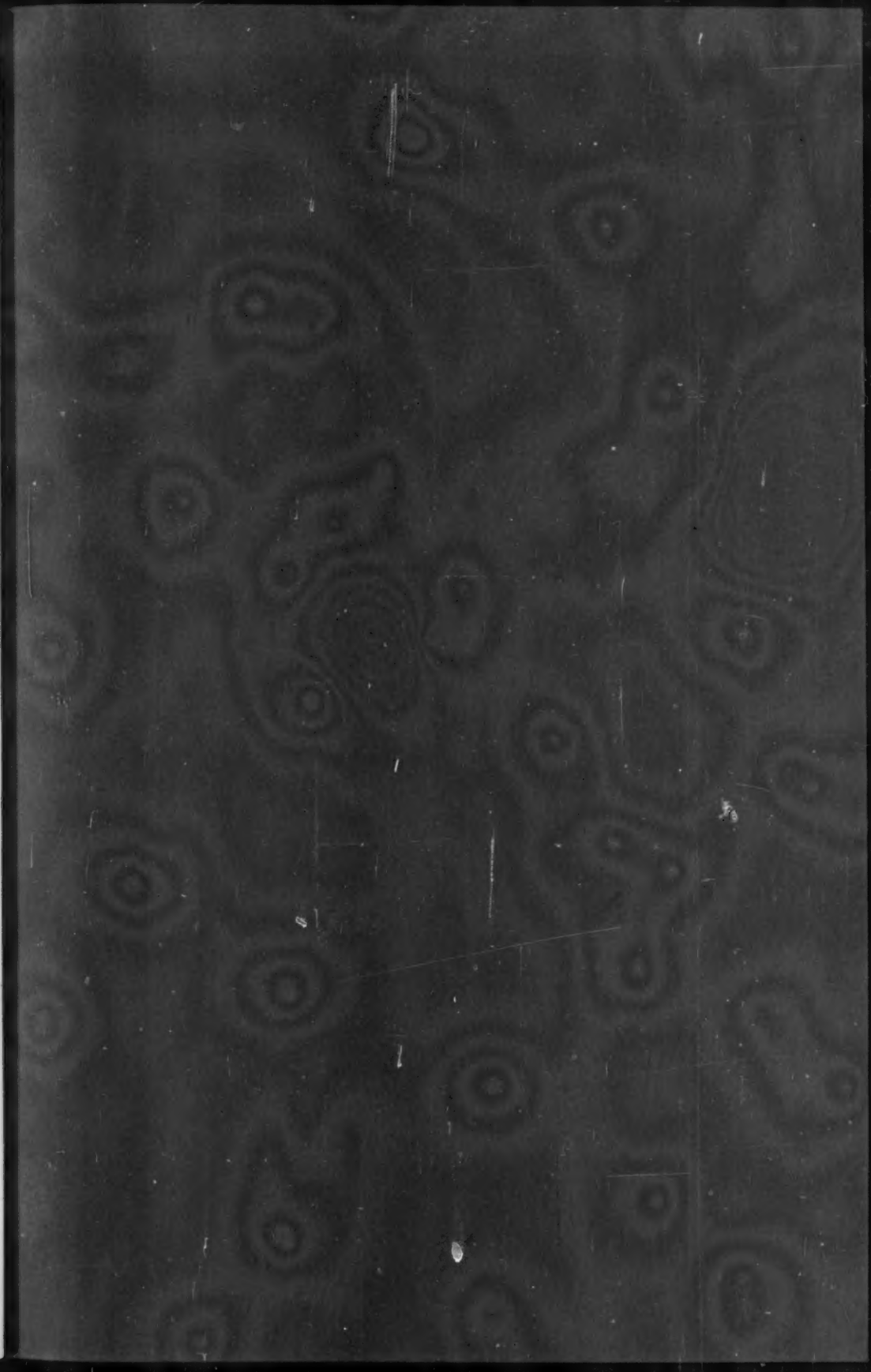
Measurement of Educational Need, The. PAUL R. MORT. No. 150, 1924.

Some Values Derived from Extensive Reading of General Science. FRANCIS DAY CURTIS. No. 163, 1924.

Study of Intelligence Test Elements, A. LEONA VINCENT. No. 152, 1924.

Tentative Inventory of the Habits of Children from Two to Four Years of Age, A. RUTH ANDRUS. No. 160, 1924.

Values of New Type Examinations in the High School. STERLING G. BRINKLEY. No. 161, 1924.



Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Morrill Hall.
Subscription $6.50. 600 pages annually. Edited by E. B. Titchener.
Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University. Subscription $7. 700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company.
Subscription $4.25. 480 pages annually.
Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company.
Subscription [illegible]20 pages annually. Psychological literature.
Monthly. [illegible] 1904. Edited by Samuel W. Fernberger.

Psychological M[illegible]as—Princeton, N. J.: Psychological Review Company.
Subscription [illegible] per vol. 500 pp. Founded 1895. Ed. by Shepherd I. Franz.
Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.: Psychological Review Company.
Subscription $1.50. 200 pp. Founded 1895. Edited by Madison Bentley.
An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116th St.
Subscription $4. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge and Wendell T. Bush.

Archives of Psychology—Sub-Station 84, N. Y.: Archives of Psychology.
Subscription $5. 500 pp. ann. Founded 1906. Ed. by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.

Journal of Abnormal Psychology and Social Psychology—Albany, N. Y.
Subscription $5. Boyd Printing and Publishing Co. Edited by Morton Prince in co-operation with Floyd H. Allport.
Quarterly. 432 pages annually. Founded 1906. Abnormal and Social.

Psychological Clinic—Philadelphia: Psychological Clinic Press.
Subscription $2.50. 288 pages. Ed. by Lightner Witmer. Founded 1907.
Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

Training School Bulletin—Vineland, N. J.: The Training School.
Subscription $1. 160 pp. ann. Ed. by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.

Comparative Psychology Monographs—Baltimore: Williams & Wilkins Co.
Subscription $5. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W.
Subscription $5. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J.
Psychological Review Company. 480 pages annually. Experimental.
Subscription $4.25. Founded 1916. Bi-monthly. Ed. by John B. Watson.

Journal of Applied Psychology—Baltimore: Williams & Wilkins Co.
Subscription $5. 400 pages annually. Founded 1917.
Quarterly. Edited by James P. Porter and William F. Book.

Journal of Comparative Psychology—Baltimore: Williams & Wilkins Co.
Subscription $5. 500 pages annually. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

Journal of Personnel Research—Baltimore: Williams & Wilkins Co.
Subscription $5.00. 500 pages approximately. Edited by Walter V. Bingham, Lewis L. Thurstone, and Clarence S. Yoakem. Monthly.